12-2016

Genetic study of carotenoids in maize grain (Zea mays L.)

Oscar Rafael Espejel Venado
Purdue University

GENETIC STUDY OF CAROTENOIDS IN MAIZE GRAIN (*Zea mays* L)

A Thesis

Submitted to the Faculty

of

Purdue University

by

Oscar Rafael Espejel Venado

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2016

Purdue University

West Lafayette, Indiana

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Oscar Rafael Espejel Venado

Entitled
Genetic Study of Carotenoids in Maize Grain (Zea mays L)

For the degree of Master of Science

Is approved by the final examining committee:

Torbert Rocheford
Chair

Mario Ferruzzi

Jianxin Ma

William Muir

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Torbert Rocheford

Approved by: Joseph Anderson 10/31/2016

Head of the Departmental Graduate Program Date

For my parents

ACKNOWLEDGEMENTS

After spending long hours studying for exams, doing research work, writing a paper and a thesis I am finally at the point where I need to say thank you to all people that believed and supported me during this adventure. One might think this is trivial compared with all those mentioned activities but the truth is other. In the next paragraphs, I want to express my feeling to every single person that deserves a recognition here.

To my parents and relatives:

Maria Elena Venado (my mom) and Jaime Espejel (my dad), who raised me with love and from a very young age they taught me important principles. Even up until now I keep learning from them and as they have told me “even if you are a street cleaner, because there is nothing bad about it or any profession, you just need to do your best. We love you and respect no matter what. You and your sister are our biggest success”. An important part of this accomplishment is also attributed to my sister Alejandra Espejel Venado, my brother in law Sergio Martinez and my two nephews Arturo and Alex Martinez Espejel. Other relatives and friend from my parents have helped and prayed for me in the past, but definitely my uncle Russ Paradise deserves also a great credit because without his helped I could not have overcome the language barrier.

To the Rocheford Lab

First of all I want to say thanks to Dr. Torbert Rocheford for accepting me in his lab and provide me with an assistantship during this time. Dr. Rocheford was not only a mentor but he was like a second father to me during my time here. Thanks to the current and previous grad students; Travis Becket, Megan Fenton, Sam Prather. Thanks to the staff associated with our lab: Marsha Kern and Jason Adams, Tyler Lawson, Sam Grisham, Dacia Garay, Lucas Krizon and Joey Stillwell for the support in the field and in the lab. Honorable mention to Dr. Brenda Owen who was like a second advisor to me and answered all my questions

To my committee member and all my professor at Purdue

Thanks to Dr. Jainxin Ma because he was willing to serve as a committee member and I enjoyed so much his class. Dr. Mario Ferruzzi that let me work in his lab and helped me to understand the aspect of food science and Dr. Bill Muir for helping with some sequence data and his assertive comments.

To my friends from high school

The so called “bolita retejaladora: Humberto Herrera, Eduardo Jiménez, Rafael León, Diana Jeronimo, Raquel Díaz, Areli Gutiérrez and Luis Dávila”. Even though we have gone through different ways and we do not see each other as often as we would like I want to say thanks you for all these years of real friendship and I wish you all the best.

To my friends from Purdue

Their invaluable friendship made my time at Purdue special and I had so many good times with you. Thank you Dan Drabiak, Zihan Wang, Patrick Kassner, Jaesik Hann, Konstantin Pandl. There is an extraordinary girl that has become a real close friend and she has a distinctive place among all my friends. A simple thank you is not enough to express my gratitude toward you Abby Pepin

To Funding sources:

Thanks to CONACyT for the scholarship 578948/409405 .With the extra money I was able to travel around USA and get to know much better the American culture

To my friend from CIMMYT

Thanks to different people from the Bioscience building: Maribel Velázquez, the wheat molecular breeding lab, the maize molecular breeding lab, the wheat quality lab and the Seed of Discovery project.

To other grad students:

The help I got from other graduate students let me accomplish my research without his help it could have been really difficult. Thanks Brad Thada (Tuinstra Lab), Darwin Ortiz and Ingrid Aragon (Ferruzzi Lab), and finally but not last Dr. Alencar Xavier (Rainy Lab).

To Purdue University:

I feel proud to be part of this great university and I do not regret about coming here to pursue a higher degree. Boiler up!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

αC	α-carotene
LUT	Lutein
βC	B-carotene
15ZβC	15-Z-β-carotene
13ZβC	13-Z-β-carotene
9ZβC	9-Z-β-carotene
βCX	β-cryptoxanthin
ZEA	Zeaxanthin
XAN	Xanthophylls
cisβC	All cis-β-carotene
proVA	Pro-Vitamin A
TC	Total Carotenoids
Ratio1	β-branch carotenoids / α-branch carotenoids
Ratio2	β-carotene / β-cryptoxanthin
Ratio3	β-carotene / (β-cryptoxanthin + Zeaxanthin)
Ratio4	β-cryptoxanthin / Zeaxanthin
Ratio5	(β-carotene + β-cryptoxanthin) / Zeaxanthin
Ratio6	β-cryptoxanthin / proVA

Ratio7	β-cryptoxanthin / TC
DOXP	1-deoxy-D-xylulose 5-phosphate synthase
IPP	isopentenyl pyrophosphate synthase
GGPPS	geranylgeranyl pyrophosphate synthase
PSY	phytoene synthase
PDS	phytoene desaturase
Z-ISO	ζ-carotene isomerase
ZDS	ζ-carotene desaturase
CRTISO	carotenoid isomerase
LCYE	lycopene ε-cyclase
LCYB	lycopene β-cyclase
CYP97A	β-carotene hydroxylase (P450)
CYP97C	ε-carotene hydroxylase (P450)
CRTRB	β-carotene hydroxylase
VDE	violaxanthin de-epoxidase
ZEP	zeaxanthin epoxidase
NCED	9-cis-epoxycarotenoid dioxygenase
CCD1	carotenoid cleavage dioxygenase 1

ABSTRACT

Espejel Venado, Oscar Rafael. M.S., Purdue University, December 2016. Genetic Study of Carotenoids in Maize Grain (*Zea mays L.*). Major Professor: Torbert Rocheford.

Pro-Vitamin A (proVA) carotenoids, which are converted into retinol (Vitamin A) in the human body, have been the subject of human nutrition studies and are a target for biofortification of staple crops. Historically, β-carotene has been the principal target for enhancing levels of proVA, yet there is recent interest in enhancing the proVA carotenoid β-cryptoxanthin. Studies have shown that β-cryptoxanthin has excellent bioavailability, and its use in maize may be nearly as effective as β-carotene in providing retinol. The primary aim of this study was to gain a better understanding of the genetic control of levels of β-cryptoxanthin, conversion of β-carotene into β-cryptoxanthin, conversion of β-cryptoxanthin into zeaxanthin, and flux into the β-branch of the carotenoid pathway. We studied a biparental population derived from inbreds with relatively high levels of β-cryptoxanthin and different ratios of β-cryptoxanthin to β-carotene. Three field replications of this $F_{2:3}$ population were grown and the grain analyzed by liquid chromatography (LC) with diode array detection. Data from a previous study using two color assessment methods, visual score and chromameter readings, were included to associate and compare LC and color results. Composite interval mapping (CIM) identified 90 quantitative trait loci (QTL). Notably, we detected a QTL for Ratio 3 (β-carotene / [β-cryptoxanthin +

zeaxanthin]) and Ratio 5 ([β-carotene + β-cryptoxanthin] / zeaxanthin) on chromosome 2 that contains candidate gene *hydroxylase 4* (*hyd4*), which has not been previously associated with QTL for carotenoids in maize grain.

A common finding for LC and chromameter data included a QTL on chromosome 5 for β-carotene, zeaxanthin, β-carotene + β-cryptoxanthin, β-branch proVA, cisβC, Total carotenoids, Ratio 1 (β-carotene / β-cryptoxanthin + zeaxanthin), h, h99 and b*. This region contains candidate gene *lycopene β cyclase* (*lcyβ*). This new information could potentially be used in biofortification breeding programs to increase the content of proVA and the total carotenoids in maize.

CHAPTER 1. QTL MAPPING USING HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC) FOR B-CRYPTOXANTHIN IN THE BIPARENTAL POPULATION HI27 X A272.

1.1 Introduction

The world's population is expected to will grow by 2.3 billion over the next 30 years. This growth is largely expected in developing countries, mainly in Sub-Saharan Africa and Latin America, which are considered regions with lower food access and high levels of undernourishment. Additionally, urban areas will expand and rural areas will shrink some, making food production a challenge in the coming years (FAO 2009). Even with a reduction in income inequality between developed and developing countries, access to a broad array of different types of food will be limited to certain regions of the world (Alexandratos and Bruinsma 2012). Hence, economic inequality and malnutrition will be major issues for a significant portion of the world's population.

Staple crops like wheat, rice and maize, which are a major source of calories and proteins (Bouis and Novenario-Reese 1997), have been produced in increasing volumes over the last 40 years due to advances in cropping methodologies, and the use of hybrids and biotechnology (Conway & Toenniessen 1999; Heinemann et al. 2013). In Sub-Saharan Africa, Latin America and some parts of Asia, these cereals are the most important food crops (Bouis and Novenario-Reese 1997). However, the amount of micronutrients such as Pro-Vitamin A (proVA) iron, and zinc are lower in these cereals compared with fruits or

green/yellow vegetables (Tang 2013). Therefore, these crops are logical targets for increasing micronutrients in their grain.

The enhancement of some micronutrients in staple crops has been accomplished by genetic engineering. However, the controversy and legal regulations surrounding genetically modified organisms have limited the use of these cultivars (Acquaah 2009). This problem is being addressed by the CGIAR (Consultative Group on International Agricultural Research) through the HarvestPlus (HP) program, which focuses on improving and proVA, iron, zinc levels through conventional breeding using natural variation (Nestel et al. 2006). Even though a number of studies have revealed insights into candidate genes affecting quantitative variation of carotenoids (Kandianis et al. 2013; Owens et al. 2014; Suwarno et al. 2015) there are still gaps in our knowledge of the genetic mechanisms underlying the biosynthetic pathway of carotenoids as relate to the proVA biofortification process in maize. In order to facilitate further progress, we need more complete knowledge of the genes involved and their regulation.

1.1.1 Carotenoids and β-cryptoxanthin

Carotenoids are naturally occurring lipid-soluble pigments with a tetraterpene structure, consisting of eight isoprenoids bound together with a general $(C_5H_8)_n$ formula (Goodwin 1980), and they differ in the presence of cycling groups or by the presence of a hydroxyl functional group (Rao and Rao 2007). Some authors mention that there are around 600 to 700 compounds (Howitt & Pogson 2006; Breithaupt et al. 2007; Rao & Rao 2007). They can also be divided into two subclasses: the first one includes those with at least one oxygen atom (O) called xanthophylls, whereas those without oxygen atoms (O)

are denominated carotenes (Breithaupt et al. 2007). The presence of these compounds confer the characteristic yellow, orange and red pigmentation in fruits and vegetables (Rao and Rao 2007).

β-cryptoxanthin (βCX) or 3R-β-β-caroten-3-ol is a natural lipid-soluble carotenoid that belongs to the xanthophylls (Goodman, 1980). βCX also resembles the structure of β-carotene (βC) but is more polar due to the presence of the hydroxyl group. It is mainly found in tropical fruits and green/red vegetables such as chili pepper, carrots, tangerines, oranges, and papaya (Burri et al. 2016)

1.1.2 Role of Carotenoids and β-cryptoxanthin

Carotenoids play an important role in plants and the function of these compounds depends on the type of tissue, which can be either photosynthetic or non-photosynthetic. The main roles of carotenoids in the photosynthetic tissue are light capture and photoprotection against oxidative stress (Yasushi 1991), whereas in the non-photosynthetic tissue the functions are different, varying from phytohormone precursors, to seed dispersal agents to pollination attractors (Bartley and Scolnik 1995; Howitt and Pogson 2006). Plants have specific structures to store these carotenoids. In the case of photosynthetic tissue, carotenoids are present in the chloroplasts with other molecules such as chlorophyll and apoproteins. In non-photosynthetic tissue carotenoids are present in structures called chromoplasts. These organelles are responsible for the coloration in flowers, fruits and some roots (Bartley and Scolnik 1995).

A main biological action of carotenoids in humans is to provide a source of antioxidants to reduce chronic diseases. There are other beneficial properties such as gap

junction communication, modulation of gene expression, cell growth regulation and modulation of drug metabolizing enzymes. In the case of β-carotene (βC), β-cryptoxanthin (βCX), and α-carotene (αC), they are precursors of vitamin A and thus are important in the prevention of some diseases associated with vitamin A deficiency (VAD) (Turner et al. 2013), such as celiac disease, cystic fibrosis, pancreatic insufficiency, duodenal bypass, chronic diarrhea, bile duct obstruction, giardiasis, and cirrhosis.

βCX, along with βC and αC, is a proVA carotenoid, but it also has other unique biological activities. βCX acts as an antioxidant in vitro (Stahl et al. 2000). the antioxidant activity of non-vitamin A-forming carotenoids is different for the vitamin A-forming carotenoids (Burri et al 2016). The proVA compounds have a complete protection against DNA damage, but the non proVA compounds have a double activity as a protectant against DNA damage or as an enhancer of DNA damage. This evidence suggests that in vivo studies are needed to determine whether βCX has this double activity. Other studies have found that βCX has a positive effect in terms of increasing bone mass by stimulating osteoblasts and inhibiting the effect of osteoclasts (Yamaguchi 2012). This effect is mediated by the expression of the insulin-like growth factor 1, the transforming growth factor β1 and the core-binding factor α1, which are proteins involved in the osteoblast differentiation process (Uchiyama and Yamaguchi 2005). Furthermore, the suppression of osteoclastogenesis is mediated by βCX inducing an apoptosis enzyme called *caspase-3* and suppressing the gene *Bcl-2*, which is an anti-apoptosis protein (Yamaguchi 2012). Moreover, it is well documented that carotenoids are highly correlated with cancer risk prevention. Matsumoto et al. (2007) showed that βCX outperformed other carotenoids in inducing the Retinol acid receptor (RAR), which many authors have found to reduce and

prevent the effect of lung cancer, emphysema and atherogenesis (Lian et al. 2006; Matsumoto et al. 2007; Liu et al. 2011), while other studies have found that this carotenoid decreased colon cancer and bladder cancer (Narisawa et al. 1999; Miyazawa et al. 2007)

1.1.3 Absortion, Bioavailability and Metabolism of β-cryptoxanthin

The absorption process depends on many factors such as food source, level of food processing, dietary factors, fat amount and fiber content in the food, food matrix and human factors (Breithaupt et al. 2007; Rao and Rao 2007; Burri et al. 2011; Goltz and Ferruzzi 2013). In general, carotenoids and β-cryptoxanthin follow a similar absorption process. This involves four steps: 1) release from the food matrix by mastication followed by enzymatic hydrolysis (Furr and Clark 1997) 2) micellization, where the carotenoids and fat-soluble compounds are solubilized (Hernell et al. 1990; Reboul 2013). 3) uptake and transfer into the intestinal epithelial cells (Goltz and Ferruzzi 2013). The uptake is carried out by two mechanisms, a) passive diffusion (Hollander and Ruble 1978), or through b) facilitative transporters like the epithelial transporter Scavenger Receptor class B type 1 (SR-B1) (van Bennekum et al. 2005), the cluster determinant 36 (CD36) (Sakudoh et al. 2010) and potentially the cholesterol membrane transporter NPC1L1 (During et al. 2005) and 4) assembly of carotenoids chylomicrons particles by a subsequent secretion into the lymphatic system (Neilson et al. 2012).

The higher bioavailability of βCX was observed in the animal model Mongolian gerbils where βCX was more effective than βC at maintaining VA status (Davis et al. 2008). Preclinical and clinical studies in humans have reported significant differences in the bioavailability of βC and βCX (Dhuique-Mayer et al. 2007; Burri et al. 2011; Turner et

al. 2013; Schweiggert et al. 2014). Thus it has been proposed that βCX might be very good source of VA. The more polar xanthophylls show more rapid micellization relative to carotenes (Fernández-García et al. 2012) and greater affinity to the SR-B1 transporter than carotenes (Kim et al. 2010). Additionally, βCX may have superior bioaccessibility compared to βC due to the hydroxyl group that confers less hydrophobicity (Burri et al, 2016). It is presently estimated that the vitamin A (VA) activity from βCX is half the VA activity from β-carotene, in that 24 mg of βCX are required to produce 1 mg of retinol, with a retinol activity equivalent (RAE) conversion ratio of 24:1. This conversion is based on the chemical structure of these carotenoids, as one molecule of β-carotene can generate two molecules of retinal, whereas βCX only generates one (Tanumihardjo 2002). Nevertheless, this RAE ratio does not seem to be calculated based on enzymatic reactions (Burri et al. 2011).

The conversion of βCX into retinol has not been completely elucidated since there are currently no studies that have specifically targeted βCX, but the available data from βC, including data from βCX, suggest that the mechanism follows the same process (Burri et al. 2016). Two main enzymes are involved: the *β-carotene-15,15' oxygenase 1* (BCO1) and the *β-carotene-9',10'-oxygenase* (BCO2). These two enzymes differ in their reaction products, as BCO1 yields a retinal by a symmetric cleavage whereas BCO2 produces long chain apocarotenoids by eccentric cleavage. Since BCO1 cleaves all proVA carotenoids, βCX is the target of this enzyme and is converted into retinal by a subsequent hydroxylation to retinol. This enzymatic reaction takes place mainly in the enterocyte but there are reports that indicate other tissues like those of the liver are also involved. There is a second

mechanism involved with BCO2, where βCX is transformed into apocarotenoids that are substrate for BCO1 (Amengual et al. 2013).

1.1.4 Genetic Studies of Carotenoids in Maize

Several authors have helped improve our understanding of the carotenoid biosynthetic pathway (Figure 1.1). Different genes are responsible for the accumulation of βCX and β-branch carotenoids. The gene, *phytoene synthase* (*psy1* or *y1*) is the first committed step is the production of phytoene from the two molecules of geranylgeranyl pyrophosphate (GGPP), thus the levels of expression of this gene can influence flux into and down the carotenoid pathway (Hirschberg 2001).

Phytoene is dehydrogenated four times and isomerized by desaturases and isomerases respectively to generate lycopene, a red-colored carotenoid (Ruiz-Sola and Rodríguez-Concepción 2012). The next step involves a branching point where lycopene is converted into carotenoids with one β ring and one ε ring such as αC and lutein (LUT) (α-branch) or carotenoids with two β rings like βC, βCX, zeaxanthin (ZEA) and other derived carotenoids (β-branch). Two genes are responsible for this branching, the *lycopene β-cyclase* (*lcyβ*) that produces the β ring carotenoids and the *lycopene ε-cyclase* (*lcyε*) that catalyzes the ε ring carotenoids. Harjes et al. (2008), using different approaches, found that variation in expression the *lcyε* locus alters the direction in the pathway towards the β-branch.

The next steps in the carotenoid pathway involve hydroxylation of cyclic carotenes. The gene *crtRB1* encodes a *β-carotene hydroxylase* responsible for converting βC into βCX. A defective allele on this locus is responsible for the accumulation of βC (Yan et al.

2010). Other hydroxylases convert βCX into ZEA; this includes the cytochrome P450 enzymes (CYP97), which work in both the β and ε rings.

Association studies have found new genes that control quantitative variation in the biosynthetic pathway and the degradation process. Owens et al. (2014) found the genes *zep1* and *lut1*. The gene *zep1* encodes the enzyme *zeaxanthin epoxidase*, responsible for the conversion of zeaxanthin into violaxanthin (Hieber et al. 2000). The *lut1*gene is a hydroxylase responsible for forming lutein. Moreover. Suwarno et al. (2015), also using association mapping, identified key genomic regions in the upstream part of the pathway, where these genes are associated with the accumulation of isoprenoid precursors such as *deoxy xylulose phosphate synthase 1* (*dxs1), geranylgeranyl pyrophosphate synthase 1* (*ggps1)* and *geranylgeranyl pyrophosphate synthase 2* (*ggps2*). This genes are important because they increased the overall flux into the carotenoid biosynthetic pathway.

New information gained on genetic control of βCX and other carotenoids may be useful to inform biofortification breeding programs for increasing proVA carotenoids. This may include selection methods that use visual or colorimeter selection and low-cost SNP assays in candidate genes for marker-assisted selection (MAS) and genomic selection in maize kernels.

1.2 Material and methods

1.2.1 Plant Material

The inbreds Hi27 and A272 were chosen for showing the highest concentrations of βCX among inbreds evaluated in the Goodman-Buckler Diversity Panel (Harjes et al. 2008). Hi27 is of subtropical background derived from the [CM104(India)*MV

source]BC6, and A272 was derived from the Boesman Open Pollinated Variety in South Africa (Liu et al. 2003). These inbreds were crossed and the F1 progeny selfed to create F2 progeny, which was selfed to create a biparental $F_{2:3}$ population with 213 individuals. The population was planted at the Agronomy Center for Research and Education (ACRE), Purdue University, West Lafayette, IN (40°28' N 86°59' W). One biological replicate was planted in 2013 and two biological replicates were planted in 2015. Plot rows were 4.5 m long with 13 $F_{2:3}$ seeds planted. Approximately seven plants were hand self-pollinated and the well-pollinated ears were harvested, shelled, bulked and stored at -80° C.

1.2.2 Carotenoid Extractions and Quantification by LC

Approximately 10 g of maize grain sampled from each of the bulks was ground to a fine powder (0.5 mm) in a Foss CT 1093 Cyclotec sample mill, and 0.6 g of each ground sample was weighed out. Carotenoid extraction and quantification was performed according to Ortiz et al. (2016). A set of 11 samples and one constant control sample ("Orange Iso", a synthetic population relatively high in total carotenoids) were extracted for each LC run. In each set, 50 µL of an internal standard, β-apo-8'-carotenal, was added to four random samples. Individual samples were mixed with 1 mL of saturated water and incubated for 10 min in an ice bath and then mixed with 5 mL of acetone. The acetone was removed into another tube and the extraction was repeated with another 5 mL of acetone. The two combined acetone phases were dried under a nitrogen evaporator. A final extraction was completed with 2 mL of methyl tert-buthyl ether and the phase was transferred into the dried tube. Another 1mL of saturated water was added and the organic phase was placed into a new tube and dried under the nitrogen stream. The samples were

resuspended in 1 mL of methanol and 1 mL of ethyl acetate. A 10µL sample was injected into a Hewlett-Packard 1090 High-performance liquid chromatographer (HPLC) with a diode array detector and a YMC C30 column (150 x 20 mm) for LC analysis. The mobile phase consisted of solvent A (Methanol: Ammonium Acetate, pH = 4.6, 98:2) and solvent B (Ethyl acetate).

The following gradient was used to separate and identify each carotenoid: 0 – 6 min 85% A to 15% B; 6 – 8 min 20% A to 80% B; 8 – 12 min 100% B and 14 – 15 min 85%A to 15% B. The detection wavelength was 450 nm and eight primary carotenoids were quantified: α- carotene (αC), lutein (LUT), β-carotene (βC), 15-z-β-carotene (15ZβC), 13-z-β-carotene (13ZβC), 9-z-β-carotene (9ZβC), β-cryptoxanthin (βCX) and zeaxanthin (ZEA). We calculated eleven carotenoid traits: six sums of primary traits, sum of βC and βCX (βC+βCX), β-branch proVA carotenoids, xanthophylls (XAN: LUT + ZEA), the cis isomers (cisβC: 15ZβC + 13ZβC + 9ZβC), proVA carotenoids (βC plus half αC, βCX, 9ZβC, 13ZβC, 15ZβC), total carotenoids (TC; all eight primary traits), and seven ratios of traits (ratios reported by other studies). Table 1.1 summarizes the ratios calculated and their biological meaning. These ratios were calculated to explain the flux in the two branches (Ratio1), conversion of a substrate into the next one or two products (Ratio2, Ratio3, Ratio4 and Ratio5) or examine relationships between a certain sets of carotenoids (Ratio6 and Ratio7). The concentration of proVA (µg g^{-1} dry matter) was calculated according to the following equation:

$$proVA = \beta C + \frac{1}{2}(\alpha C + \beta CX + \ cis\text{-}\beta\text{-}carotenes)$$

This assumes that βC can be cleaved and that it produces two molecules of retinol; whereas with αC, βCX and the cisβC isomers produce only one molecule of retinol, thereby providing only half the activity of β-carotene (Burri et al. 2011; Ortiz et al. 2016).

1.2.3 Genotyping and Map Construction

From the $F_{2:3}$ population, tissue samples from ten plants per genotype were bulked, lyophilized, and shipped to DuPont Pioneer where DNA was extracted and genotyped with the 766 chip Illumina SNP markers. Monomorphic markers and makers showing segregation distortion were excluded and only 231 polymorphic markers were used to create a genetic map with the Rqtl package (Broman et al. 2003), as described by Owens 2015.

1.2.4 Statistical Analysis

The grand mean for the three replicates for each plot was used in QTL analysis for carotenoid concentration. QTL analysis was performed with composite interval mapping (CIM) using the software WinQTL cartographer version 2.5 (Basten et al. 1994). Permutation thresholds (Churchill and Doerge 1994) were estimated for the purpose of assessing the significance of the results obtained in CIM. A stepwise regression procedure with $p < 0.05$ as significance level was used and a window size of 10 cM with a walking speed set of 2 cM was used. The B73 RefGen v2 positions of the nearest flanking SNP markers to the QTL were used to annotate the candidate genes present in the QTL interval.

1.3 Results

1.3.1 Carotenoid Concentration

For the 21 carotenoid traits evaluated, 15 showed a skewed distribution (Figure 1.2) and others did not (15βC, βCX, ZEA, [βC+βCX], XAN and TC), with non-normality confirmed with the Shapiro test ($p = 0.05$). The carotenoid values from each replicate and the averages for the three replicates are presented in Table 1.2. The total carotenoid concentration over the three replicates, averaged between 21.93 and 48.86 µg/g dry weight (DW) and the concentration of proVA ranged from 3.01 to 11.9 µg/g DW. The mean for β-cryptoxanthin was 4.23 µg/g DW, with a range of 1.72 – 6.67 µg/g DW, and contributed 36.3 to 58.6 % of the proVA concentration per family. This relationship is consistent with the high Pearson correlation between βCX and proVA concentration ($r = 0.70$), which is to be expected as βCX is a component of proVA. Moderate correlations were observed for βC with βCX and βCX with ZEA, 0.47 and 0.54, respectively, consistent with BC being the precursor of BCX and BCX the precursor of ZEA. Low correlations were observed for the relationships of βCX with LUT and αC consistent with previous reports. Moderate to high correlations were observed for traits with a product-substrate relationship in the carotenoid biosynthetic pathway, such as βC to βCX, or βCX to ZEA, similar to other previous reports (Owens et al. 2014; Suwarno et al. 2015). Traits in different branches of the pathway showed lower correlations.

Heritabilities for the primary traits, sums of traits and ratios are presented in Table 1.3. Overall the heritabilities were relatively high for all the traits ranging from 0.58 to 0.84 with the exception of αC with 0.32, consistent with previous reports. A ratio may be more

consistent across environments than absolute values of the individual traits that comprise the ratio; consequently they could be expected to have higher heritabilities.

1.3.2 Linkage Map

The genetic linkage map consisted of 231 polymorphic markers across the ten chromosomes. The length of the linkage groups ranges from 96.9 centiMorgan (cM) for chromosome 9 to 244.2 cM in chromosome 1 (Figure 1.4). The average distance between markers was 6.5 cM.

1.3.3 QTL Mapping for Carotenoids

Numerous QTL were associated in the biparental population Hi27×A272 with variation in the concentrations of primary carotenoids, sums of carotenoids, and ratios of carotenoids. A total of 90 QTL were identified (Table 1.4), covering all chromosomes except 4. Ratio 6 (βCX / proVA) showed the highest number of QTL, with nine, followed by LUT, ZEA, Ratio 2 (βC / βCX), Ratio 4 (βCX / ZEA) and Ratio 7 (βCX / TC) with six. The smallest number of QTL was identified for αC with just one. The number of QTL mapped for other traits ranged from two to five (Table 1.4).

A number of QTL for different traits were co-located at the same region. QTL for βCX and ratios involving βCX were associated with peak SNP PHM15961.13 on chromosome 6. QTL for the traits LUT, 13ZBC, β-branch proVA, XAN, cisβC, TC and Ratio 6 (βCX / proVA) overlapped the same region on chromosome 1. Other overlapping results were found on chromosome 7 and 8. QTL for the traits αC, βC, 15ZβC, 13ZβC, ZEA, sum of βC + βCX, β-branch proVA carotenoids, XAN, cisβC, proVA, TC and Ratio1

mapped to an overlapping region on chromosome 5 from 57.8 to 80.8 cM. These QTL were highly significant, with LOD values ranging from 5.27 to 12.38. The maximum phenotypic variance explained (PVE) by a QTL in this region was 18.2% for ZEA

For the 90 QTL identified, 31 contained underlying candidate genes, of which 14 were associated with one than more trait. The gene *phytoene synthase* (*y1*, GRMZM2G300348) is involved in the first committed step in the carotenoid biosynthetic pathway, and was associated with Ratio3 (βC / [βCX + ZEA]), Ratio 6 (βCX / proVA) and Ratio 7 (βCX / TC). Other relatively upstream pathway candidate genes (Figure 1.1) were associated with carotenoid QTL: *zeta-carotene desaturase1* (*zds1*, GRMZM2G454952) with LUT, 15ZβC and Ratio 6 (βCX / proVA); and *carotene isomerase 2* (*crti2,* GRMZM2G106531) with Ratio 3 (βC / [βCX + ZEA]) and Ratio 5 ([βC + βCX] / ZEA). The carotenoid biosynthetic pathway splits into the α-branch and β-branch through actions of *lycopene epsilon cyclase* (*lcyE,* GRMZM2G012966) and *lycopene β cyclase* (*lcyB*, GRMZM5G849107). The gene *lcyE* lies within the QTL region associated with ZEA, XAN, Ratio 2 (βC / βCX), Ratio 4 (βCX / ZEA) and Ratio 5; and the gene *lcyB* lies within a QTL region with β-branch related traits: βC, 13ZβC, ZEA, sum of βC + βCX, β-branch proVA carotenoids, XAN, cisβC, TC and Ratio 1 (β-branch / α-branch). The LOD and PVE values for the QTL associated with *lcyB* were among the highest in the study (Table 1.4).

Downstream candidate genes in the β-branch of the pathway were associated with QTL: *β-carotene hydroxylase 1* (*crtRB1* or *hyd3,* GRMZM2G164318) for Ratio 3 (βC / [βCX + ZEA]), and *hydroxylase 4* (*hyd4,* GRMZM2G164318) for Ratio 3 and Ratio 5 ([βC + βCX] / ZEA). Since both Ratio 3 and Ratio 5 are similar relationships involving βC, βCX

and ZEA (See Table 1.1) it is not surprising to find QTL with intervals that contain candidate genes that control the flux of βC, βCX and ZEA. The gene *cytochrome P450 13* (*cyp13*, GRMZM5G837869) was associated with QTL for ZEA and TC.

Four genes involved in the catabolism of carotenoids were located within QTL confidence intervals. The *9-cis-epoxycarotenoid dioxygenase 1* gene (*nced1*, GRMZM2G014392) for LUT, βC, 13ZβC, β-branch proVA carotenoids, XAN, cisβC, TC. The *9-cis-epoxycarotenoid dioxygenase 4* gene (*nced4,* GRMZM2G408158) and *carotenoid cleavage dioxygenase 7* (*ccd7,* GRMZM2G408158) in the same interval on chromosome 2 were linked with Ratio 3 (βC / [βCX +ZEA]) and Ratio 5 ([βC + βCX] / ZEA). The *9-cis-epoxycarotenoid dioxygenase 8* gene (*nced8,* GRMZM5G838285) on chromosome 5 was associated with XAN and TC.

Three candidate genes from the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway, which is involved in the production of the isoprenoid precursors of carotenoids, were located within a large confidence interval of a QTL on chromosome 5 for ZEA, XAN and TC. The genes *geranylgeranyl hydrogenase1* (*ggh1*, GRMZM2G105644), *hydroxymethylbutyl diphosphate synthase 1* (*hds1*, GRMZM2G137409) and *methyl erythritol cyclodiphosphate synthase 2* (*mecs2*, AC209374.4_FG002) lie within this interval framed by the flanking markers of peak SNP PHM1870.20.

In order to better visualize and understand the effect of some carotenoid candidate genes we plotted marker class means of specific carotenoid traits. A relationship between marker class means and the allele composition was clear for *hyd4* (Figure 1.6). When the marker class means of other genes were plotted, some of the relationships were less clear (Figures 1.7 and 1.8). Marker class means were plotted to search for combined effects for

two genes. An almost significant interaction between *hyd4* and *y1* was observed (Figure 1.9), whereas other interactions were not close to significant (data not shown). More research in larger populations is needed to better understand and make conclusions about the way genes interact to affect the carotenoid concentrations.

1.3.4 QTL Mapping from Colorimeter Data

A QTL analysis on chromameter readings (colorimeter) and visual scores was previously performed (Owens 2015) in the biparental population Hi27xA272. Detailed information on the methods, results and discussion can be found in the Owens 2015 study. Briefly, two methods were used to measure color in the 2013 replicate. Visual color scores were assigned based on the color, with 1 representing the lightest color (yellow) and 9 representing the darkest orange color. Colorimeter values were determined by using the Konica Minolta CR400 chromameter. Colorimeter values on three scales were used: the Hunter scale (L*, a* and b*) and the Commission Internationale d'Eclairage (CIE) scale (L, a and b) are based on the Opponent-Color Theory. The Konica Minolta (http://www.hunterlab.com/an-1005b.pdf), scale recorded the values L99, a99, b99, c99 and h99 which are also based on the same theory but with some modification by the manufacturer.

A total of 71 QTL were identified for color measurements: 5 QTL for visual score, and 66 QTL for the three scales of colorimeter. The PVE for visual score ranged from 6.01 to 13.81%, and the PVE for colorimeter ranged from 5.03 to 28.19%. The three scales combined (CIE, Hunter and the Konica Minolta) identified 8 co-located QTL. The QTL regions, for colorimeter and visual score, were spread across the chromosomes 1, 3, 4, 5,

6, 9 and 10. Common results between the visual score and the color scale were found on two regions. On chromosome 9 the traits visual score, h, h99, L*, a*, b*, L, b, L99, a99 and b99 identified the same QTL with support interval 61.4-72.6 cM whereas on chromosome 10 the traits visual score, h, h99, L*, L, b, L99, and b99 were associated with the same QTL. Table 1.5 summarizes the results for the QTL associated with color.

Comparison of the candidate genes associated with LC based carotenoid values and the colorimeter values in a single replicate from 2013 revealed the *lcyB* and the *nced1* as common findings. The traits associated with *lcyB* were βC, ZEA, sum of βC and βCX, β-branch proVA carotenoids, XAN, cisβC, TC, Ratio 1 (β-branch / α-branch), h, h99 and b* with PVE that range from 7.19-18.2% in a high majority of the traits the LOD values were high. The QTL on chromosome 1 that contained the *nced1* was found on traits LUT, 13ZβC, β-branch proVA, cisβC, TC and Visual score. The PVE ranged from 2.68-6.01 %. Table 1.6 includes all the genes found independently in both studies.

1.4 Discussion

Different approaches and genetic materials have been used previously to better elucidate the genetic architecture of carotenoids in maize grain. In this study, a biparental population, Hi27xA272, that has a range of carotenoid concentrations and intensity of kernel color was used. We identified new candidate genes not previously associated with LC-determined levels of carotenoids in maize endosperm. The study from Owens 2015 detected new associations using colorimeter and visual color score, including multiple isoprenoid pathway, carotenoid pathway and carotenoid degradation genes. All approaches

provided support to some previously reported candidates genes associated with levels of carotenoids.

New information was revealed in this study on genetic control of βCX levels and associated ratios. A single QTL was detected for βCX. Owens et al. (2014) in a genome wide study (GWAS) of a small maize association panel reported a significant SNP close to this QTL region on chromosome 7 for βCX. Phenotypic variation for the traits ZEA, Ratio 1 (β-branch / α-branch) and Ratio 6 (βCX / proVA) were also associated with this QTL. A candidate gene associated with the biosynthetic pathway or degradation of carotenoids was not identified in this interval. There may be sequence variation at this QTL that acts in a regulatory manner, as shown by other studies these types of regions can control levels of carotenoids (Welsch et al. 2007; Lopez et al. 2008; Toledo-Ortiz et al. 2010).

Carotenoid metabolic candidate genes were found in QTL regions detected for Ratio 5 ([βC + βCX] / ZEA). We identified a QTL whose support interval includes *hyd4*, a logical candidate gene in the conversion of βC to βCX and βCX to ZEA. Suwarno et al. (2015), using a genome-wide analysis in a carotenoid association mapping panel, reported associations with SNPs in the coding regions of *hyd1* and *hyd5* with levels of βCX. Sun et al. (1996) suggested that *hyd5* encodes a hydroxylase responsible for the conversion of βCX into ZEA in *Arabidopsis thaliana*. Li et al. (2010) cloned and characterized a cDNA (ZmBCH1) encoding a β-carotene hydroxylase with 99.4% identity with the amino sequence with *hyd4*. They found that ZmBCH1 was able to convert β-carotene into β-cryptoxanthin and zeaxanthin in an *in vitro* assay. Furthermore, *hyd4* and *crtRB1(hyd3)* share high homology (Vallabhaneni et al. 2009) and have a conserved hydroxylase domain with a plastid-targeting signal described by Sun et al. (1996). These reports support the

hypothesis that *hyd4* encodes a functional hydroxylase capable of metabolizing both β-carotene and β-cryptoxanthin, consistent with our derived Ratio 5 ([βC+βCX] / ZEA) for which we detected a QTL in this region. Gene variations in the promotor, the untranslated or intronic regions might be responsible for the specificity of substrate and efficiency in the hydroxylation reaction (Yan et al. 2010) and contribute to the differences between both hydroxylases *crtRB1* and *hyd4*.

The chromosome 2 interval at 99.5-109.7 cM for Ratio 5 ([βC+βCX] / ZEA) also contained other candidate genes, *nced4*, *ccd7* and *crti2*. The degradation genes *nced4* and *ccd7* could potentially be significant for this trait since they cleave βC (Vallabhaneni et al. 2010). Pan et al. (2016) demonstrated that *ccd7* plays an important role in the stringolactone biosynthesis with specific activity in the root compared with other tissues like leaf, ear or tassel, and the *nced4* gene has been implicated in plant responses to elevated temperature (Huo et al. 2013). Yet neither *nced4* nor *ccd7* are highly expressed in maize endosperm tissue in inbred B73 (Sekhon et al. 2011). The *crti2* gene is an upstream carotenoid biosynthetic pathway gene that encodes a redox-type enzyme that provides isomers needed to produce lycopene (Li et al. 2007), and since we did not identify *crti2* in other QTL it is not highly likely to affect downstream compounds. These genes were detected within a large support interval in our low resolution map, and we hypothesize that *hyd4*, with its capacity to convert βC and βCX into ZEA, is reasonably a more likely candidate for influencing Ratio 5. This result indicates that further investigation is needed to determine the roles of the various hydroxylases in regulating relative carotenoid concentrations.

We studied the conversion of βC into subsequent products through Ratio 2 (βC / βCX), and Ratio 3 (βC / [βCX+ZEA]). A QTL detected for Ratio 2 mapped to a QTL region containing the candidate gene *lycE*, which was also detected for Ratio 4 (βCX / ZEA). The *lycE* gene influences the distribution of carotenoids between the α-branch and the β-branch by catalyzing the formation of the αC precursor δ-carotene from lycopene (Figure 1.1). This further supports the importance of this QTL region for the accumulation of β-branch carotenoids since an allele of *lcyE* with reduced function leads to increase the flux into the β-branch (Harjes et al. 2008). We would expect all of the β-branch compounds to be affected similarly, however the effect plots for the peak SNP PZA03005.19 (Figure 1.8), did not provide a clear indication of the way this gene influences levels of carotenoids. Zeaxanthin and β-carotene values showed the opposite trend to β-cryptoxanthin with respect to this SNP. This observation supports the idea that a form of feedback inhibition affects levels of compounds within the β-branch. Further information is needed to elucidate the relationship of the various carotenoid hydroxylases and their possible interactions with the lycopene cyclases.

Ratio 3 (βC / [βCX+ZEA]), is a somewhat similar biologically to Ratio 5 ([βC+βCX] / ZEA); therefore it was not surprising to find the same QTL region involved for both Ratio 3 and Ratio 5. Two other QTL were mapped for Ratio 3; one on chromosome 6 containing the candidate gene *phytoene synthase* (*y1*) and the other on chromosome 10, with peak SNP PZA01456.2 lying 236 kbp upstream of *crtRB1* or *hyd3*. *Phytoene synthase* has been associated with variation in carotenoid levels in tomatoes (Thorup et al. 2000) and the *crtRB1* gene has been associated with variation in maize, and it is the target of biofortification breeding programs (Yan et al. 2010; Babu et al. 2013; Dhliwayo et al.

2014). Fu et al. (2013) described that a combined effect of the most favorable haplotypes for *y1*, *lcyE* and *crtRB1* increased the levels of βC 10.5 fold in maize, in contrast to the study from Babu et al. (2013) which showed that when individual favorable haplotypes are together in the same line, an associated decrease the in concentration of βC occurred. The QTL detected for the colorimeter traits did not contain *y1* or *lycE*, which have often been associated with carotenoid traits.

Other hydroxylases besides *crtRB1*and *hyd4* were detected such as *cyp13, cyp15, hyd7* and *hyd8* in both colorimeter and carotenoid traits. Maize appears to have a relatively high number of β-ring hydroxylase genes, and while it is unclear if they have overlapping activity, it is very likely that they play an important role in carotenoid composition and turnover. This concept is supported in the results presented by Tian et al. (2003). They described the implication of β-hydroxylase 1 and 2 single and double knockout mutants in vivo and concluded that: a) carotenoid hydroxylases have overlapping activity, b) their activity partially compensates, and c) there are other hydroxylases that are not characterized in the pathway. Furthermore, the tissue, gene expression pattern and pathway regulation must be considered to properly identify the specific function of all these hydroxylases.

Our results further support the concept that a subset of different carotenoid pathway biosynthetic genes with favorable alleles could significantly increase the concentration of carotenoids in maize endosperm. Current efforts have reached the value of 15 µg/g DW initial goal of biofortification breeding programs (Pixley et al. 2013). This has been accomplished by selecting favorable variation in the genes *y1*, *lycE* and *crtRB1* which have been proposed to increase the flux into the overall pathway, flux into the β-branch, and

limit conversion of beta-carotene to beta-cryptoxantin and zeaxanthin, respectively, and consequently result in an increase in levels of proVA carotenoids in the endosperm (Li et al. 2008; Babu et al. 2013; Fu et al. 2013). Since *y1*, *crtRB1*, and *hyd4* were present in regions underlying QTL for Ratio 5 ([βC+βCX] / ZEA), we propose that *hyd4* may be a member of the subset of genes contributing to useful variation in proVA content. Hence, biofortification breeding programs could potentially select favorable alleles for *hyd4* to increase the content of β-carotene and β-cryptoxanthin. This idea is supported by the effect plot (Figure 1.6) where one parental allele at SNP PZA02080-1 (peak SNP for the QTL underlying *hyd4)* was associated with a higher level of proVA concentration. Different combination of haplotypes at the *lcyE* and *crtRB1* loci were stated to potentially increase the content of βC up to 12 fold (Yan et al. 2010), and the inclusion of additional genes such as *hyd4* may also be involved.

Different physicochemical factors as well as genetic factors are responsible for postharvest degradation, which can account for a reduction of 40-70% in the concentration of proVA over 6 months of storage (Mugode et al. 2014). Two families of genes have been indicated in the catabolism of carotenoids. One group is the carotenoid cleavage dioxygenases (ccd), selective oxidative cleavage non-hemi iron (II) dependent enzymes (Harrison and Bugg 2014). The second group is the 9-cis epoxycarotenoid dioxygenases (nced), tissue-specific β-ione ring cleaving enzymes (Tan et al. 2003; Vallabhaneni et al. 2009) and crucial in the synthesis of abscisic acid (ABA) (Schwartz et al. 1997). Our study located QTL support intervals containing genes associated with these families. A QTL near *ccd1* and different *nced* genes (*nced1*, *nced3*, *nced9*) were associated with colorimeter data. Ahmed et al. (2002) used a combination of color scales to address thermal degradation of

carotenoids and change in color in papaya, thus making colorimeter a suitable tool for evaluating degradation of carotenoids and detecting quantitative variations useful in QTL analysis. Independently, we associated colorimeter and LC based values with genes in the nced family. Notably the nced genes were mainly linked with levels of β-branch carotenoids. This is probably because ZEA, the last product in the β-branch, is the main precursor in the ABA biosynthetic pathway and some of the nced genes cleave specific β-branch carotenoids (Xiong and Zhu 2003). We also identified association of these genes with some derived traits: TC, XAN, Ratio 3, Ratio 5 and Ratio 6 (βCX / proVA), with the exception of Ratio 6 the other carotenoid traits are related with ZEA and in the case of the ratios with the flux into the β-branch. Thus, both LC and color methods provide insights about associations with degradation genes of carotenoids, which can be explored further.

Lycopene beta cyclase (*lcyB*) is a key enzyme in the synthesis of proVA carotenoids in the β-branch as it converts lycopene into β-carotene (Singh et al. 2003). Chander et al. (2008) identified a QTL on chromosome 5 for ZEA but they did not explicitly state *lcyB* was a candidate gene. Rather Chander et al. (2008) proposed that the QTL on chromosome 5 is only encountered in populations containing specific genetic backgrounds (Wong et al. 2004; Chander et al. 2008). In our study some traits were associated with a QTL region containing the *lcyB* metabolic gene. The high LOD scores at the peak positions presented here suggest a strong association with quantitative variation in the content of β-branch carotenoids and with color. This QTL region on chromosome 5 was also associated with traits related to the β-branch of the carotenoid pathway. Furthermore, lcy*β* was used in the genetic engineering of "Golden Rice" (Ye et al. 2000) and the high carotenoid maize line M37W-Ph4 (Zhu et al. 2008) mainly to increase the β-branch carotenoids. Additionally,

the detection of common QTL for multiple traits involving β-branch compounds has been previously described (Wong et al. 2004; Chander et al. 2008; Harjes et al. 2008; Kandianis et al. 2013).

In this study QTL for colorimeter traits h, h99 and b* mapped to the *lcyB* region. The relationship of the QTL with these traits is not as evident as the LC-based carotenoid traits due to the absence of product-substrate association. However, the b*, b and b99 traits in the color scale measure the yellow hues which are an important component of orange color. Moreover, the correlation between the a*, a and a99 traits and β-carotene is among the highest between colorimeter and LC values. This suggests that a combination of color values might be more informative than a single one. The Meléndez-Martínez et al. (2007) study on oranges support this idea, since the β-branch carotenoids showed particular values in the a*, b* plane in contrast with other carotenoids e.g. α-branch or reddish carotenoids (lycopene). Therefore, we propose that colorimeter can be used to detect QTL regions containing carotenoid biosynthetic genes in the β-branch. Additionally, the overlapping results between LC and colorimeter based analysis suggests that colorimeter may be a quick breeding tool to evaluate for relative concentrations of β-branch carotenoids, and notable β-carotene, during early stages of selection where large number of individuals are inexpensively screened.

Upstream genes are important because they influence the metabolic flux and exhibit pleiotropic effects with downstream genes (Clotault et al. 2012). Thus, we included in our study upstream genes mainly in the methylerthritol 4-phosphate (MEP) pathway, which provides the precursor geranylgeranyl diphosphate (GGPP), the first substrate in the carotenoid biosynthetic pathway. The genes *dxr1*, *dxs1* and *dxs2* have been associated with

classical color loci (Andorf et al. 2010). Remarkably, these genes have been studied in yellow by white crosses whereas here we associate them in a yellow by orange cross. Consequently, these genes are logical candidates for color QTL since they have shown an effect o0n kernel color.

1.5 Summary

In this study, liquid chromatography was used to identify chromosomal regions associated with quantitative variation of carotenoids. In these results, and the previous results from Owens (2015), a number of carotenoid candidate genes, and also isoprenoid genes were identified. Notably, *lcyβ* and *hyd4* are of particular interest because: a) *lcyβ* is a node gene that controls the flux into the β-branch, a defective or very weak allele will likely result in a lower content of proVA and this can negatively impact the production of ABA and b) *hyd4* is likely to affect the content of βCX, and evidence suggests this carotenoid is a potential source of proVA, thus unlocking the possibility to look for favorable alleles that can increase the content of this carotenoid. Our evidence suggests that *hyd4* might be as important a target to consider as *crtRB1*. Further studies are needed to characterize the allelic variation at these loci and the development of user-friendly molecular markers to distinguish the effective alleles that will be necessary to make this information more useful to breeders.

1.6 Literature Cited

Acquaah G (2009) Principles of plant genetics and breeding. John Wiley & Sons

Ahmed J, Shivhare US, Sandhu KS (2002) Thermal degradation kinetics of carotenoids and visual color of papaya puree. J Food Sci 67:2692–2695.

Alexandratos N, Bruinsma J (2012) World agriculture towards 2030/2050: the 2012 revision. ESA Working paper Rome, FAO

Amengual J, Widjaja-Adhi MAK, Rodriguez-Santiago S, et al (2013) Two carotenoid oxygenases contribute to mammalian provitamin A metabolism. J Biol Chem 288:34081–34096.

Andorf CM, Lawrence CJ, Harper LC, et al (2010) The Locus Lookup tool at MaizeGDB: identification of genomic regions in maize by integrating sequence information with physical and genetic maps. Bioinformatics 26:434–436.

Babu R, Rojas NP, Gao S, et al (2013) Validation of the effects of molecular marker polymorphisms in LcyE and CrtRB1 on provitamin A concentrations for 26 tropical maize populations. Theor Appl Genet 126:389–399.

Bartley GE, Scolnik PA (1995) Plant carotenoids: pigments for photoprotection, visual attraction, and human health. Plant Cell 7:1027.

Basten CJ, Weir BS, Zeng Z (1994) Zmap–a QTL cartographer. In: In. Citeseer,

Bouis HE, Novenario-Reese MJG (1997) The determinants of demand for micronutrients. International Food Policy Research Institute (IFPRI)

Breithaupt DE, Yahia EM, Velázquez FJV (2007) Comparison of the absorption efficiency of α-and β-cryptoxanthin in female Wistar rats. Br J Nutr 97:329–336.

Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890.

Burri BJ, Chang JST, Neidlinger TR (2011) β-Cryptoxanthin-and α-carotene-rich foods have greater apparent bioavailability than β-carotene-rich foods in Western diets. Br J Nutr 105:212–219.

Burri BJ, La Frano MR, Zhu C (2016) Absorption, metabolism, and functions of β-cryptoxanthin.

Chander S, Guo YQ, Yang XH, et al (2008) Using molecular markers to identify two major loci controlling carotenoid contents in maize grain. Theor Appl Genet 116:223–233.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Clotault J, Peltier D, Soufflet-Freslon V, et al (2012) Differential selection on carotenoid biosynthesis genes as a function of gene position in the metabolic pathway: a study on the carrot and dicots. PLoS One 7:e38724.

Conway G, Toenniessen G (1999) Feeding the world in the twenty-first century. Nature 402:C55–C58.

Davis C, Jing H, Howe JA, et al (2008) β-Cryptoxanthin from supplements or carotenoid-enhanced maize maintains liver vitamin A in Mongolian gerbils (Meriones unguiculatus) better than or equal to β-carotene supplements. Br J Nutr 100:786–793.

Dhliwayo T, Palacios-Rojas N, Crossa J, Pixley K V (2014) Effects of S Recurrent Selection for Provitamin A Carotenoid Content for Three Open-Pollinated Maize Cultivars. Crop Sci 54:2449–2460.

Dhuique-Mayer C, Borel P, Reboul E, et al (2007) β-Cryptoxanthin from citrus juices: assessment of bioaccessibility using an in vitro digestion/Caco-2 cell culture model. Br J Nutr 97:883–890.

During A, Dawson HD, Harrison EH (2005) Carotenoid transport is decreased and expression of the lipid transporters SR-BI, NPC1L1, and ABCA1 is downregulated in Caco-2 cells treated with ezetimibe. J Nutr 135:2305–2312.

FAO UN (2009) How to Feed the World in 2050. In: Rome: High-Level Expert Forum.

Fernández-García E, Carvajal-Lérida I, Jarén-Galán M, et al (2012) Carotenoids bioavailability from foods: From plant pigments to efficient biological activities. Food Res Int 46:438–450.

Fu Z, Chai Y, Zhou Y, et al (2013) Natural variation in the sequence of PSY1 and frequency of favorable polymorphisms among tropical and temperate maize germplasm. Theor Appl Genet 126:923–935.

Furr HC, Clark RM (1997) Intestinal absorption and tissue distribution of carotenoids. J Nutr Biochem 8:364–377.

Goltz SR, Ferruzzi MG (2013) Carotenoid bioavailability: influence of dietary lipid and fiber. In: Carotenoids and human health. Springer, pp 111–128

Goodwin TW (1980) Nature and distribution of carotenoids. Food Chem 5:3–13.

Harjes CE, Rocheford TR, Bai L, et al (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science (80-) 319:330–333.

Harrison PJ, Bugg TDH (2014) Enzymology of the carotenoid cleavage dioxygenases: reaction mechanisms, inhibition and biochemical roles. Arch Biochem Biophys 544:105–111.

Heinemann JA, Agapito-Tenfen SZ, Carman JA (2013) A comparative evaluation of the regulation of GM crops or products containing dsRNA and suggested improvements to risk assessments. Environ Int 55:43–55.

Hernell O, Staggers JE, Carey MC (1990) Physical-chemical behavior of dietary and biliary lipids during intestinal digestion and absorption. 2. Phase analysis and aggregation states of luminal lipids during duodenal fat digestion in healthy adult human beings. Biochemistry 29:2041–2056.

Hieber AD, Bugos RC, Yamamoto HY (2000) Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. Biochim Biophys Acta (BBA)-Protein Struct Mol Enzymol 1482:84–91.

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol 4:210–218.

Hollander D, Ruble PE (1978) Beta-carotene intestinal absorption: bile, fatty acid, pH, and flow rate effects on transport. Am J Physiol Liver Physiol 235:G686–G691.

Howitt CA, Pogson BJ (2006) Carotenoid accumulation and function in seeds and non-green tissues. Plant Cell Environ 29:435–445.

Huo H, Dahal P, Kunusoth K, et al (2013) Expression of 9-cis-EPOXYCAROTENOID DIOXYGENASE4 is essential for thermoinhibition of lettuce seed germination but not for seed development or stress tolerance. Plant Cell 25:884–900.

Kandianis CB, Stevens R, Liu W, et al (2013) Genetic architecture controlling variation in grain carotenoid composition and concentrations in two maize populations. Theor Appl Genet 126:2879–2895.

Kim Y-S, Park C-S, Oh D-K (2010) Hydrophobicity of residue 108 specifically affects the affinity of human β-carotene 15, 15′-monooxygenase for substrates with two ionone rings. Biotechnol Lett 32:847–853.

Li F, Murillo C, Wurtzel ET (2007) Maize Y9 encodes a product essential for 15-cis-ζ-carotene isomerization. Plant Physiol 144:1181–1189.

Li F, Vallabhaneni R, Yu J, et al (2008) The maize phytoene synthase gene family: overlapping roles for carotenogenesis in endosperm, photomorphogenesis, and thermal stress tolerance. Plant Physiol 147:1334–1346.

Li Q, Farre G, Naqvi S, et al (2010) Cloning and functional characterization of the maize carotenoid isomerase and β-carotene hydroxylase genes and their regulation during endosperm maturation. Transgenic Res 19:1053–1068.

Lian F, Hu K, Russell RM, Wang X (2006) β-Cryptoxanthin suppresses the growth of immortalized human bronchial epithelial cells and non-small-cell lung cancer cells and up-regulates retinoic acid receptor β expression. Int J cancer 119:2084–2089.

Liu C, Bronson RT, Russell RM, Wang X-D (2011) β-Cryptoxanthin supplementation prevents cigarette smoke-induced lung inflammation, oxidative damage, and squamous metaplasia in ferrets. Cancer Prev Res 4:1255–1266.

Liu K, Goodman M, Muse S, et al (2003) Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. Genetics 165:2117–2128.

Lopez AB, Van Eck J, Conlin BJ, et al (2008) Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. J Exp Bot 59:213–223.

Matsumoto A, Mizukami H, Mizuno S, et al (2007) β-Cryptoxanthin, a novel natural RAR ligand, induces ATP-binding cassette transporters in macrophages. Biochem Pharmacol 74:256–264.

Miyazawa K, Miyamoto S, Suzuki R, et al (2007) Dietary b-cryptoxanthin inhibits N-butyl-N-(4-hydroxybutyl) nitrosamine-induced urinary bladder carcinogenesis in male ICR mice. Oncol Rep 17:297–304.

Mugode L, Ha B, Kaunda A, et al (2014) Carotenoid retention of biofortified provitamin a maize (Zea mays L.) after Zambian traditional methods of milling, cooking and storage. J Agric Food Chem 62:6317–6325.

Narisawa T, Fukaura Y, Oshima S, et al (1999) Chemoprevention by the Oxygenated Carotenoid β-Cryptoxanthin of N-Methylnitrosourea-induced Colon Carcinogenesis in F344 Rats. Japanese J cancer Res 90:1061–1065.

Neilson A, Ferruzzi M, Coulston AM, Boushey CJ (2012) Bioavailability and metabolism of bioactive compounds from foods. Nutr Prev Treat Dis 407–423.

Nestel P, Bouis HE, Meenakshi J V, Pfeiffer W (2006) Biofortification of staple food crops. J Nutr 136:1064–1067.

Ortiz D, Rocheford T, Ferruzzi MG (2016) Influence of Temperature and Humidity on the Stability of Carotenoids in Biofortified Maize (Zea mays L.) Genotypes during Controlled Postharvest Storage. J Agric Food Chem 64:2727–2736.

Owens BF (2015) Association and QTL Mapping of Carotenoid and Color Loci in Maize Grains.

Owens BF, Lipka AE, Magallanes-Lundback M, et al (2014) A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels. Genetics 198:1699–1716.

Pan X, Zheng H, Zhao J, et al (2016) ZmCCD7/ZpCCD7 encodes a carotenoid cleavage dioxygenase mediating shoot branching. Planta 1–12.

Pixley K, Rojas NP, Babu R, et al (2013) Biofortification of maize with provitamin A carotenoids. In: Carotenoids and Human Health. Springer, pp 271–292

Rao AV, Rao LG (2007) Carotenoids and human health. Pharmacol Res 55:207–216.

Reboul E (2013) Absorption of vitamin A and carotenoids by the enterocyte: focus on transport proteins. Nutrients 5:3563–3581.

Ruiz-Sola MÁ, Rodríguez-Concepción M (2012) Carotenoid biosynthesis in Arabidopsis: a colorful pathway. Arab B e0158.

Sakudoh T, Iizuka T, Narukawa J, et al (2010) A CD36-related transmembrane protein is coordinated with an intracellular lipid-binding protein in selective carotenoid transport for cocoon coloration. J Biol Chem 285:7739–7751.

Schwartz SH, Tan BC, Gage DA, et al (1997) Specific oxidative cleavage of carotenoids by VP14 of maize. Science (80-) 276:1872–1874.

Schweiggert RM, Kopec RE, Villalobos-Gutierrez MG, et al (2014) Carotenoids are more bioavailable from papaya than from tomato and carrot in humans: a randomised cross-over study. Br J Nutr 111:490–498.

Sekhon RS, Lin H, Childs KL, et al (2011) Genome-wide atlas of transcription during maize development. Plant J 66:553–563.

Singh M, Lewis PE, Hardeman K, et al (2003) Activator mutagenesis of the pink scutellum1/viviparous7 locus of maize. Plant Cell 15:874–884.

Stahl W, Heinrich U, Jungmann H, et al (2000) Carotenoids and carotenoids plus vitamin E protect against ultraviolet light–induced erythema in humans. Am J Clin Nutr 71:795–798.

Sun Z, Gantt E, Cunningham FX (1996) Cloning and functional analysis of the β-carotene hydroxylase of Arabidopsis thaliana. J Biol Chem 271:24349–24352.

Suwarno WB, Pixley K V, Palacios-Rojas N, et al (2015) Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. Theor Appl Genet 128:851–864.

Tan B, Joseph LM, Deng W, et al (2003) Molecular characterization of the Arabidopsis 9-cis epoxycarotenoid dioxygenase gene family. Plant J 35:44–56.

Tang G (2013) Using Plant Foods Rich in β-Carotene to Combat Vitamin A Deficiency. 13 Sight Life 20.

Tanumihardjo SA (2002) Factors influencing the conversion of carotenoids to retinol: bioavailability to bioconversion to bioefficacy. Int J Vitam Nutr Res 72:40–45.

Thorup TA, Tanyolac B, Livingstone KD, et al (2000) Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc Natl Acad Sci 97:11192–11197.

Tian L, Magallanes-Lundback M, Musetti V, DellaPenna D (2003) Functional analysis of β-and ε-ring carotenoid hydroxylases in Arabidopsis. Plant Cell 15:1320–1332.

Toledo-Ortiz G, Huq E, Rodríguez-Concepción M (2010) Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. Proc Natl Acad Sci 107:11626–11631.

Turner T, Burri BJ, Jamil KM, Jamil M (2013) The effects of daily consumption of β-cryptoxanthin–rich tangerines and β-carotene–rich sweet potatoes on vitamin A and carotenoid concentrations in plasma and breast milk of Bangladeshi women with low vitamin A status in a randomized controlled trial. Am J Clin Nutr 98:1200–1208.

Uchiyama S, Yamaguchi M (2005) b-Cryptoxanthin stimulates cell proliferation and transcriptional activity in osteoblastic MC3T3-E1 cells. Int J Mol Med 15:675–682.

Vallabhaneni R, Bradbury LMT, Wurtzel ET (2010) The carotenoid dioxygenase gene family in maize, sorghum, and rice. Arch Biochem Biophys 504:104–111.

Vallabhaneni R, Gallagher CE, Licciardello N, et al (2009) Metabolite sorting of a germplasm collection reveals the hydroxylase3 locus as a new target for maize provitamin A biofortification. Plant Physiol 151:1635–1645.

van Bennekum A, Werder M, Thuahnai ST, et al (2005) Class B scavenger receptor-mediated intestinal absorption of dietary β-carotene and cholesterol. Biochemistry 44:4517–4525.

Welsch R, Maass D, Voegel T, et al (2007) Transcription factor RAP2. 2 and its interacting partner SINAT2: stable elements in the carotenogenesis of Arabidopsis leaves. Plant Physiol 145:1073–1085.

Wong JC, Lambert RJ, Wurtzel ET, Rocheford TR (2004) QTL and candidate genes phytoene synthase and ζ-carotene desaturase associated with the accumulation of carotenoids in maize. Theor Appl Genet 108:349–359.

Xiong L, Zhu J-K (2003) Regulation of abscisic acid biosynthesis. Plant Physiol 133:29–36.

Yamaguchi M (2012) Role of carotenoid-cryptoxanthin in bone homeostasis.

Yan J, Kandianis CB, Harjes CE, et al (2010) Rare genetic variation at Zea mays crtRB1 increases [beta]-carotene in maize grain. Nat Genet 42:322–327.

Yasushi K (1991) New trends in photobiology: structures and functions of carotenoids in photosynthetic systems. J Photochem Photobiol B Biol 9:265–280.

Ye X, Al-Babili S, Klöti A, et al (2000) Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science (80-) 287:303–305.

Zhu C, Naqvi S, Breitenbach J, et al (2008) Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. Proc Natl Acad Sci 105:18232–18237.

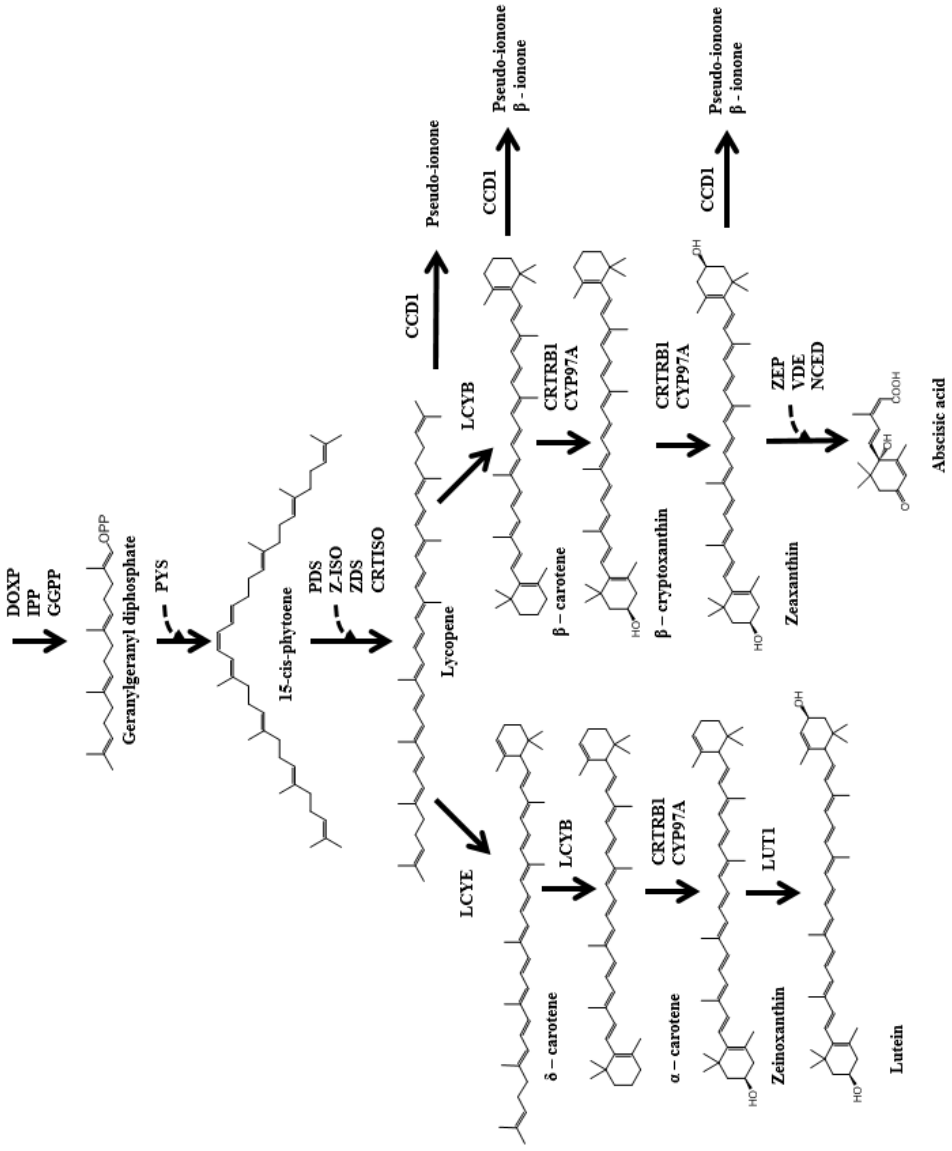


Figure 1.1 Carotenoid biosynthetic pathway

Table 1.1 Summary of Ratios and their biological meaning

Name	Mathematical relationship	Biological meaning
Ratio1	β-branch carotenoids:α-branch carotenoids [βC + βCX + ZEA + CisβC] / [αC + LUT]	Flux between the carotenoids in the β-branch versus the α-branch
Ratio2	βC / βCX	Conversion of βC into the next product βCX
Ratio3	βC / [βCX + ZEA]	Conversion of βC into the next two product βCX and ZEA
Ratio4	βCX / ZEA	Conversion of βCX into the next product ZEA
Ratio5	[βC + βCX] / ZEA	Conversion of the two main proVA carotenoids from the β-branch into ZEA
Ratio6	βCX / proVA	Contribution of βCX to the concentration of proVA
Ratio7	βCX / TC	Contribution of βCX to the total concentration of carotenoids

Abbreviations: Lutein (LUT), zeaxanthin (ZEA), β-cryptoxanthin (βCX), α-carotene (αC), β-carotene (βC) and Cis-carotenes (CisβC), pro Vitamin A (proVA) and total carotenoids (TC).

Table 1.2 Carotenoid concentration from three different replicates and the average of all of them in the biparental population Hi27 x A272.

Trait	13WLN			15WLN-A			15WLN-B			3 Replicates		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
α-Carotene	0.516	0.186	0.11 - 1.97	0.539	0.201	0.15 - 1.45	0.508	0.209	0.17 - 1.56	0.520	0.131	0.20 - 1.03
lutein	4.227	0.929	2.21 - 10.51	4.960	1.316	1.64 - 14.18	4.628	1.276	2.16 - 12.5	4.613	1.052	2.65 - 13.34
β-Carotene	2.259	0.352	0.99 - 3.24	4.05	1.315	1.25 - 8.56	3.922	1.170	0.17 - 7.87	3.399	0.794	0.15 - 6.09
15-Z-β-carotene	1.388	0.231	0.6 - 2.1	1.463	0.527	0.22 - 3.45	1.270	0.428	0.39 - 2.87	1.372	0.277	0.8 - 2.18
13-Z-β-carotene	0.740	0.238	0.32 - 3.22	0.993	0.289	0.33 - 2.23	0.862	0.266	0.13 - 2.36	0.864	0.194	0.45 - 1.80
9-Z-β-carotene	0.985	0.213	0.47 - 1.79	1.404	0.444	0.39 - 2.74	1.246	0.375	0.58 - 3.21	1.210	0.260	0.65 - 2.64
β-Cryptoxanthin	3.343	0.627	0.58 - 5.13	5.003	1.296	1.82 - 9.19	4.376	0.995	1.62 - 7.34	4.227	0.840	1.72 - 6.67
zeaxanthin	14.251	2.696	5.82 - 22.04	20.216	4.832	8.24 - 41.85	18.464	3.638	8.98 - 27.55	17.57	3.193	9.67 - 27.61
xanthophylls	18.443	3.127	9.14 - 27.19	25.176	5.49	9.88 - 49.25	23.091	4.260	11.46 - 33.53	22.275	3.614	13.99 - 33.40
cis-β-carotenes	3.115	0.547	2.02 - 5.80	3.861	1.023	1.6 - 7.37	3.379	0.797	1.36 - 7.62	3.446	0.577	2.01 - 5.29
pro-Vitamin A	5.747	0.780	3.82 - 8.04	8.754	2.179	3.45 - 14.9	8.053	1.763	3.01 - 13.33	7.497	1.291	4.74 - 11.39
Total Carotenoids	27.678	4.141	15.4 - 40.43	38.632	8.144	15.38 - 67.91	35.274	6.280	16.24 - 51.97	33.780	5.176	21.93 - 48.86
Ratio1: β-branch/α-branch	4.940	0.790	1.02 - 6.95	6.172	1.207	1.42 - 9.93	6.069	1.167	1.59 - 8.91	5.722	1.206	1.02 - 9.93
Ratio2: βC/βCX	0.702	0.227	0.23 - 3.29	0.828	0.255	0.406 - 2.34	0.921	0.293	0.03 - 2.83	0.818	0.228	0.38 - 2.58
Ratio3: βC/(βCX + ZEA)	0.131	0.023	0.05 - 0.28	0.160	0.039	0.094 - 0.366	0.172	0.042	0.01 – 0.41	0.154	0.030	0.08 – 0.38
Ratio4: βCX/ZEA	0.238	0.042	0.099 - 0.42	0.251	0.055	0.137 - 0.486	0.24	0.049	0.13 - 0.45	0.243	0.043	0.13 - 0.44
Ratio5: (βC + βCX)/ZEA	0.400	0.054	0.27 - 0.63	0.453	0.083	0.274 - 0.767	0.453	0.077	0.28 - 0.78	0.439	0.060	0.31 - 0.72
Ratio6: βCX/proVA	0.581	0.082	0.15 - 0.85	0.578	0.098	0.277 - 0.827	0.550	0.098	0.19 -0.96	0.570	0.082	0.21 - 0.79
Ratio7: βCX/TC	0.121	0.017	0.02 - 0.18	0.130	0.023	0.051 - 0.207	0.124	0.022	0.04 - 0.18	0.125	0.019	0.04 - 0.19

Eight carotenoids where measure α-carotene, lutein, β-carotene, 15-Z-β-carotene, 13-Z-β-carotene, 9-Z-β-carotene, β-cryptoxanthin, zeaxanthin. Four derived traits were calculated: xanthophylls, cis-β-carotenes, pro-Vitamin A (proVA) and total carotenoids (TC). Additionally seven ratios where calculated the Ratio1 β-branch/α-branch; Ratio2 conversion βC/βCX; Ratio3 βC / (βCX + ZEA); Ratio4 conversion βCX / ZEA; Ratio5 conversion (βC + βCX)/ZEA, Ratio6 conversion βCX/proVA and Ratio7 conversion βCX/TC

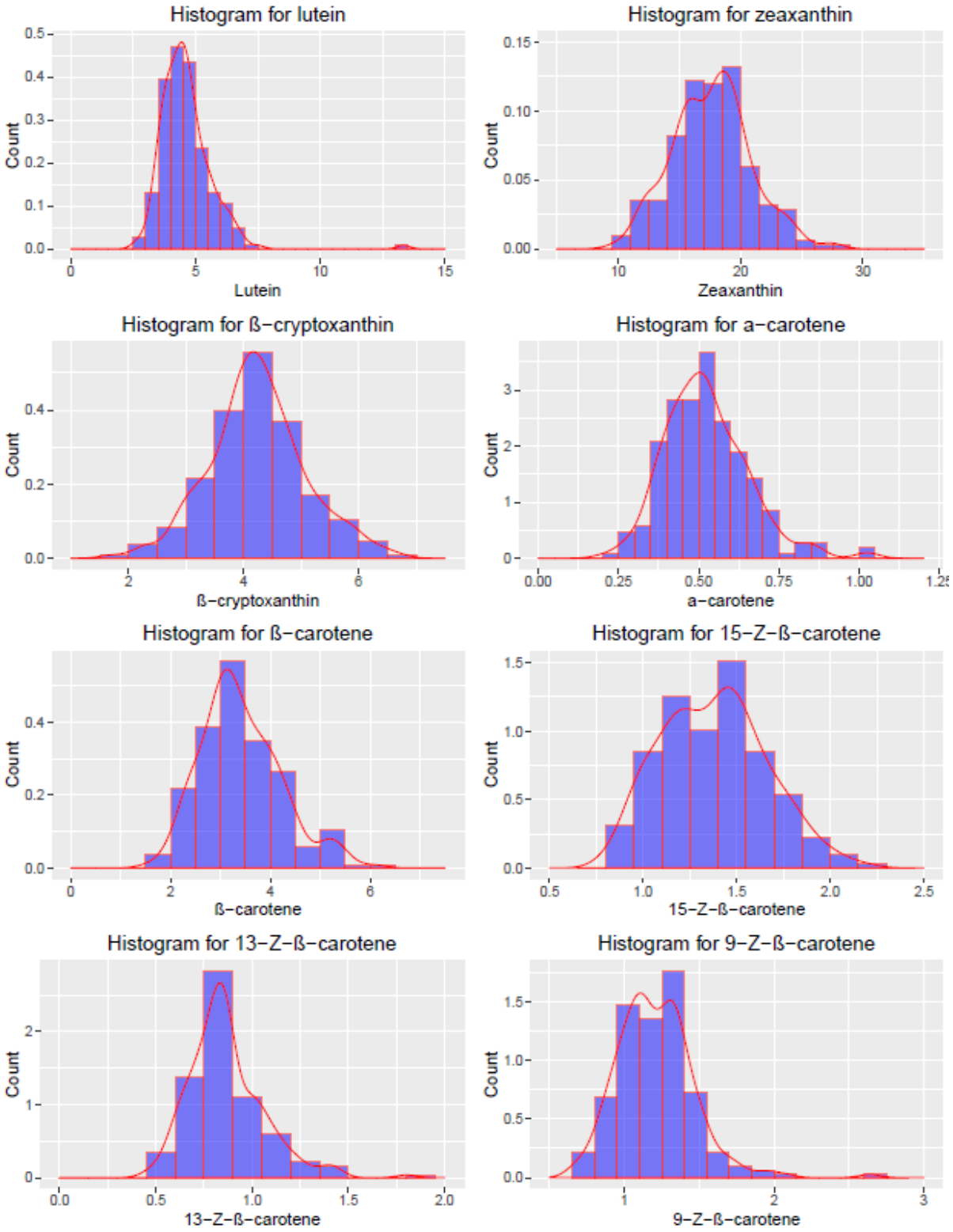


Figure 1.2 Distribution for all the carotenoids traits in the biparental population Hi27 x A272.

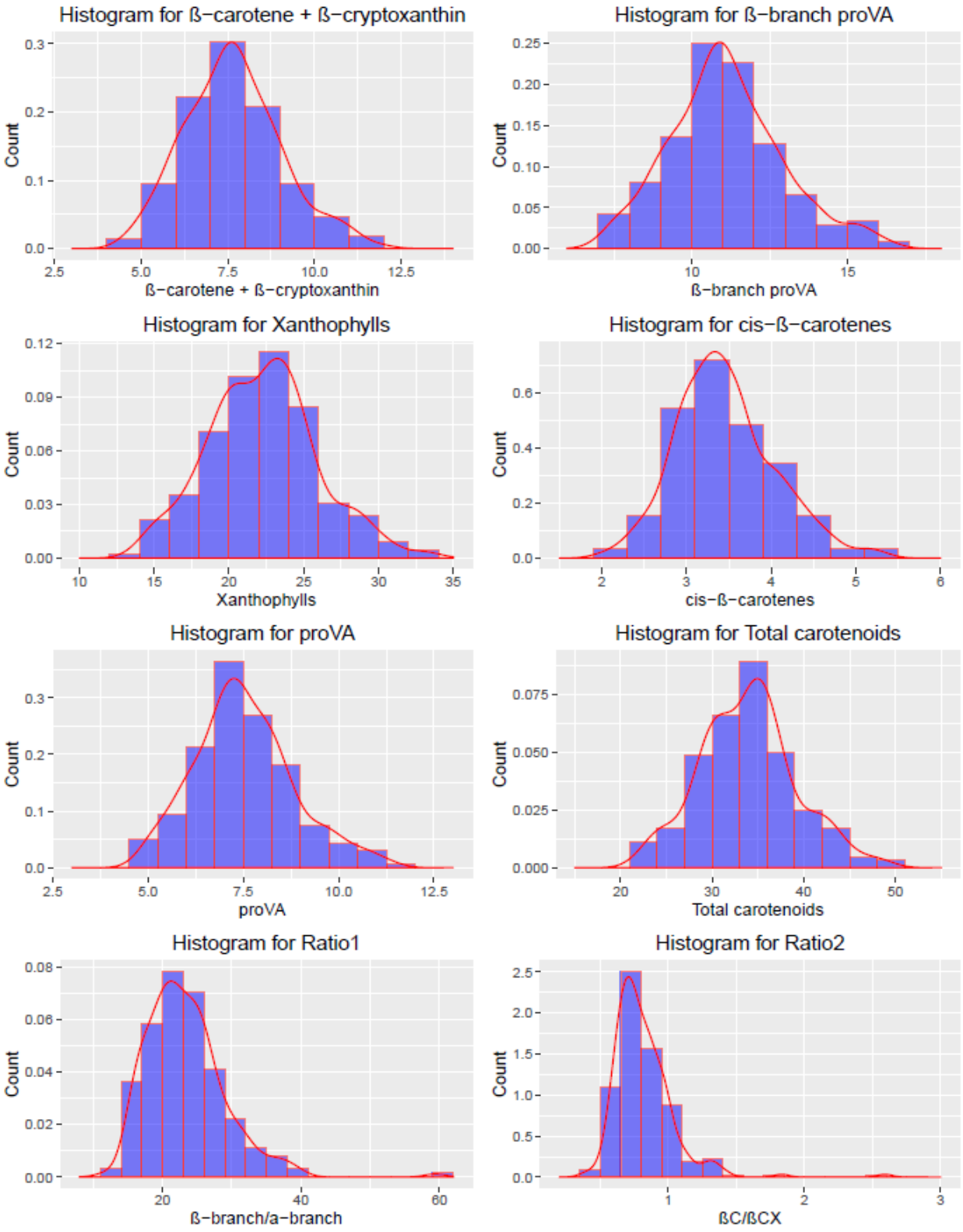


Figure 1.2 Continued. Distribution for all the carotenoids traits in the biparental population Hi27 x A272.

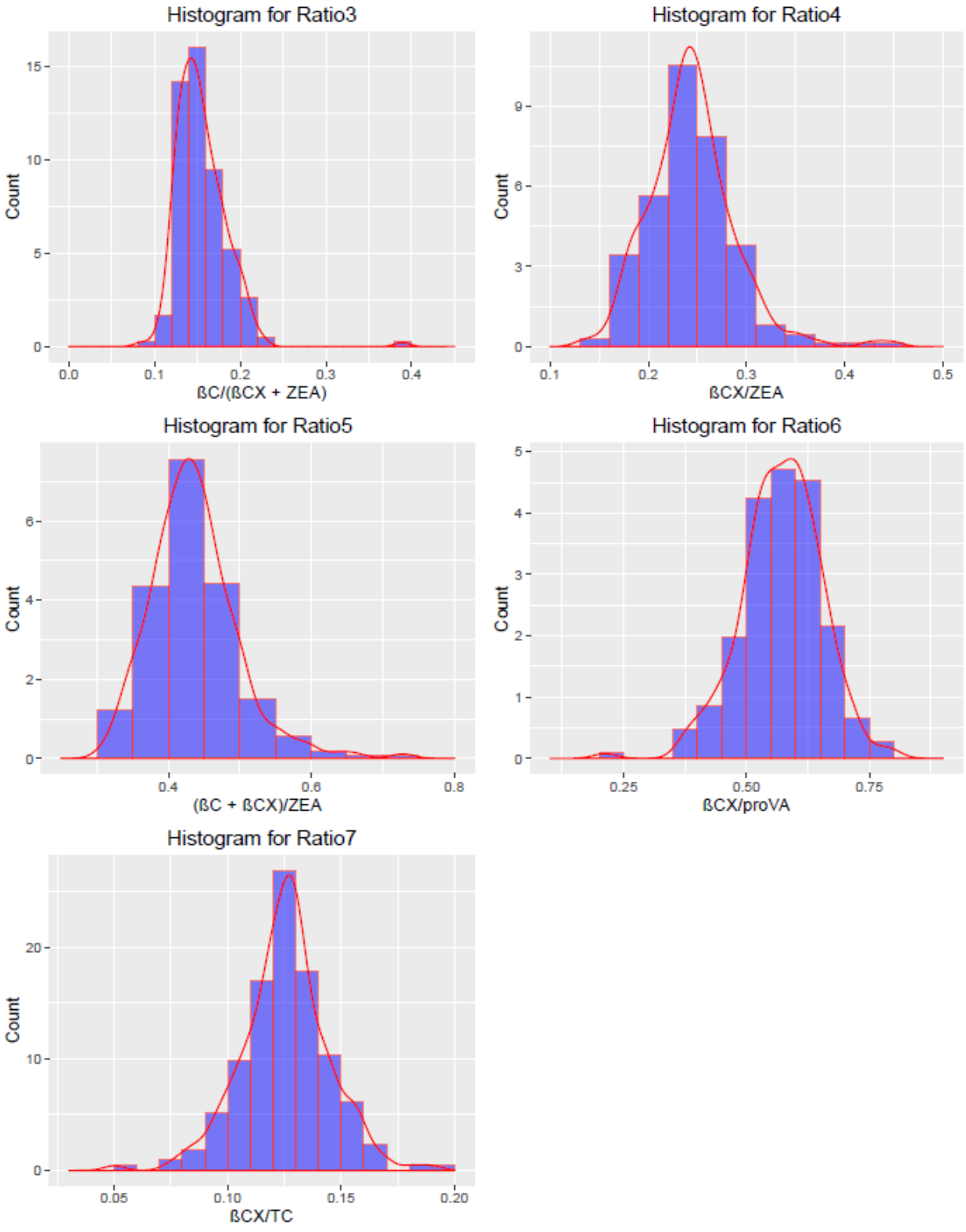


Figure 1.2 Continued. Distribution for all the carotenoids traits in the biparental population Hi27 x A272.

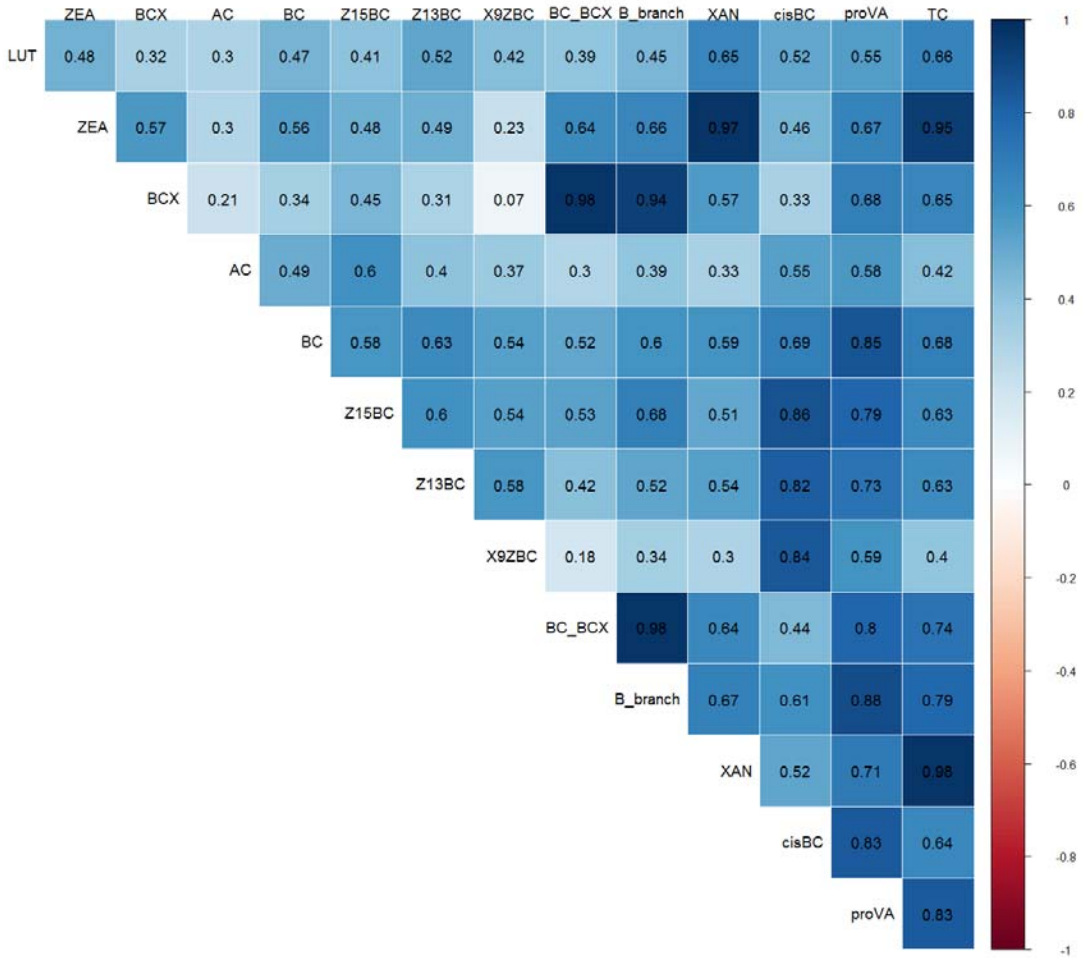


Figure 1.3 Pearson's correlation between carotenoids traits in the biparental population Hi27 x A272. Direct traits and derived traits are coded as followed LUT (lutein), ZEA (zeaxanthin), BCX (β-cryptoxanthin), AC (α-carotene), BC (β-carotene), Z15BC (15-Z-β-carotene), Z13C (13-Z-β-carotene), Z9BC (9-Z-β-carotene), BC_BCX (β-carotene + β-cryptoxanthin), B_branch (β-branch), XAN (Xanthophylls), cisBC (Cis-β-carotenes), TC (total carotenoids) and proVA (pro-Vitamin A).

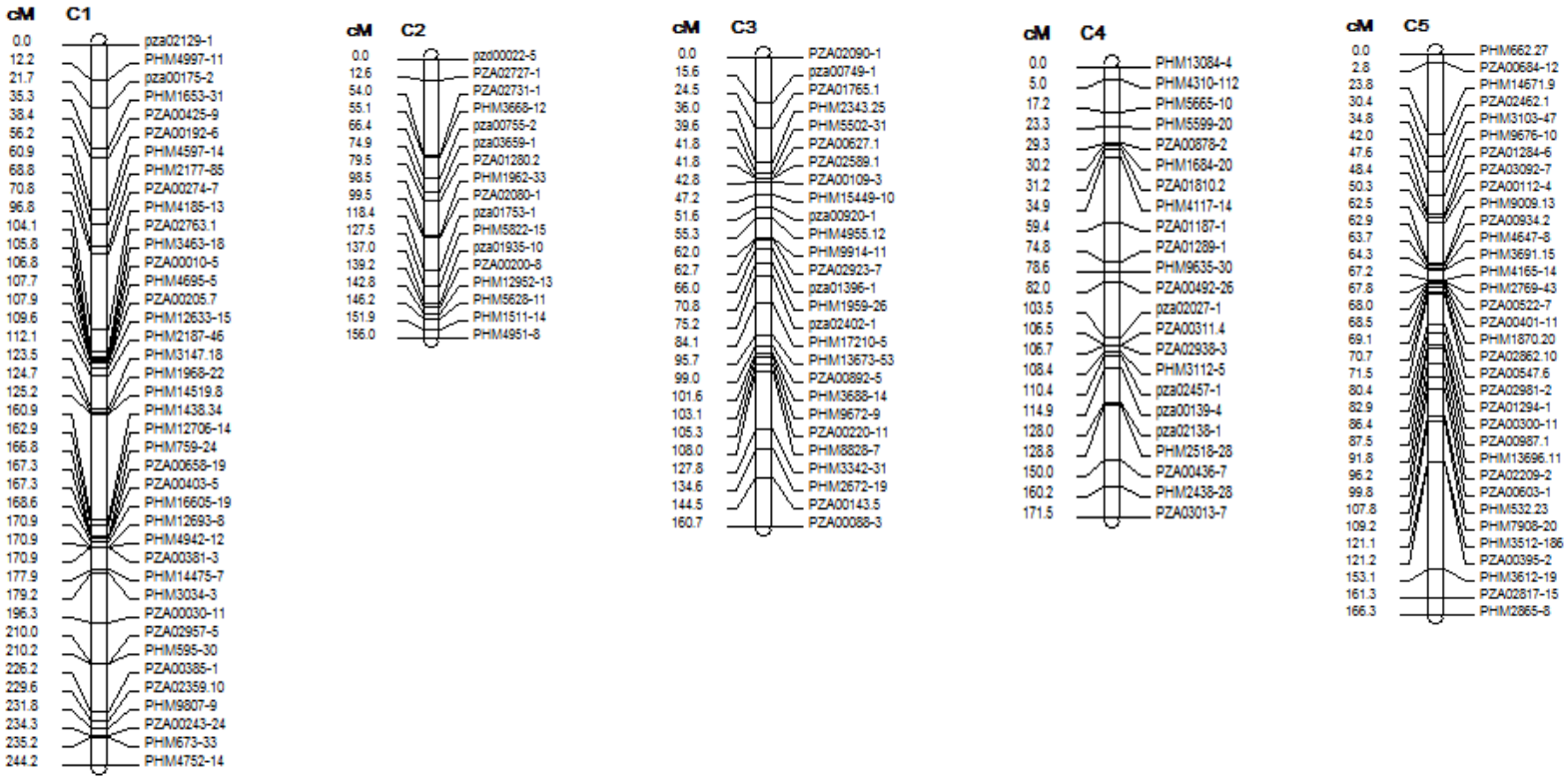


Figure 1.4 Genetic map, for the biparental population Hi27xA272 in centiMorgans (cM). The chromosome 1 has 40 markers with a distance of 244.2 cM; chromosome 2 has 17 markers, distance of 156.0 cM; chromosome 3 has 27 markers, distance of 160.7 cM; chromosome 4 has 23 markers, distance of 171.5 cM; chromosome 5 has 34 markers, distance of 166.3 cM; chromosome 6 has 21 markers, distance of 99.0 cM; chromosome 7 has 16 markers, distance of 97.7 cM; chromosome 8 has 22 markers, distance of 172.5 cM; chromosome 9 has 13 markers, distance of 96.9 cM; finally chromosome 10 has 18 markers with a distance of 131.1 cM.

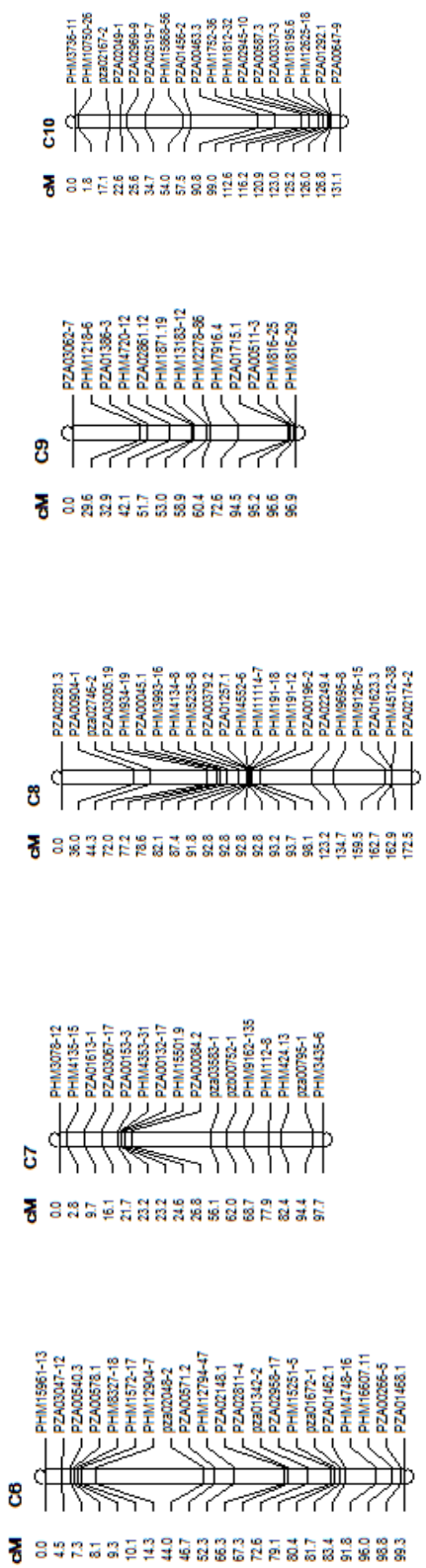


Figure 1.4 Continued. Genetic map, for the biparental population Hi27xA272 in centiMorgans (cM)

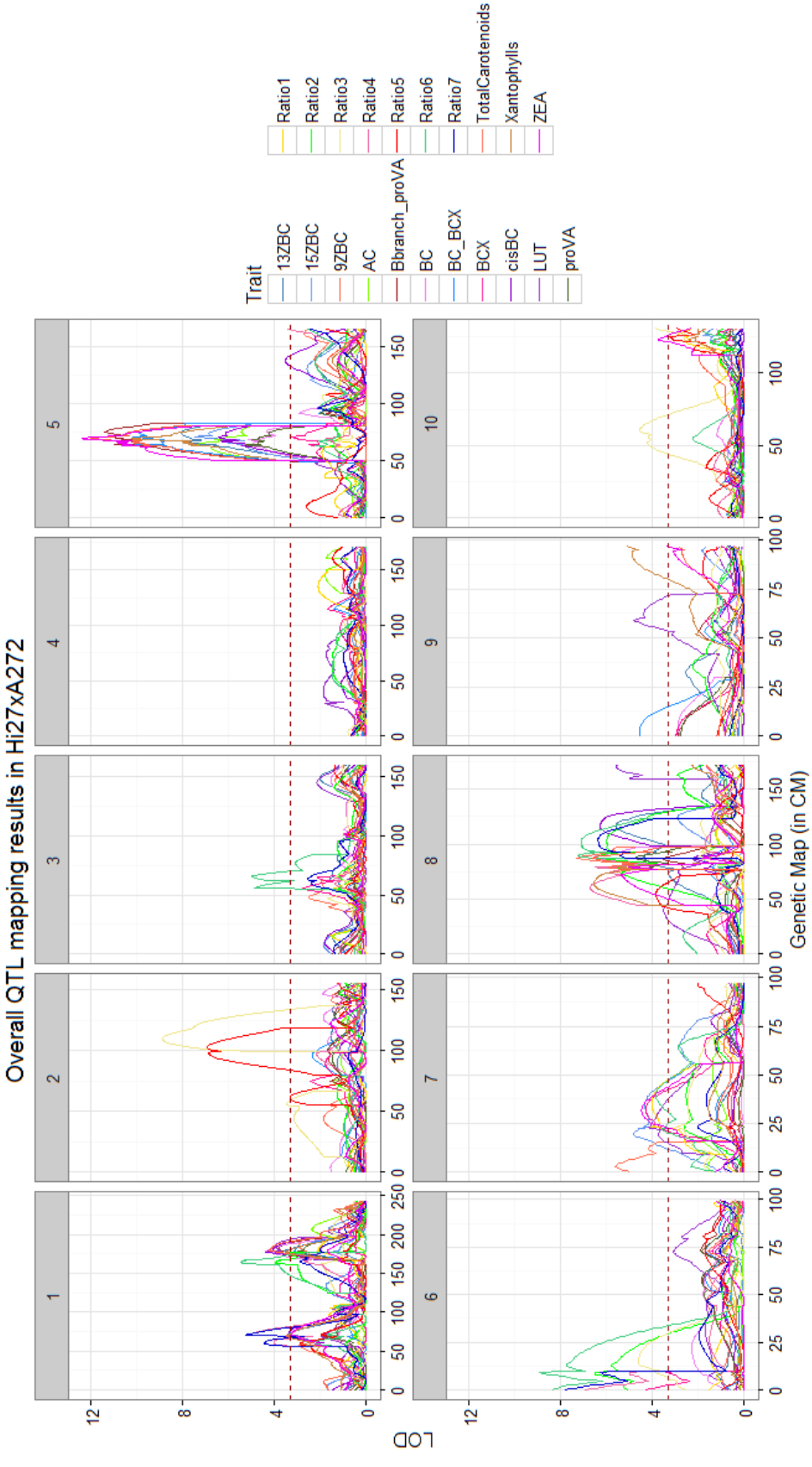


Figure 1.5 Continued. Genetic map, for the biparental population Hi27xA272 in centiMorgans (cM)

Table 1.3 Heritability of different carotenoids traits

Trait	Heritability (h^2)
α-carotene	0.325
lutein	0.701
β-Carotene	0.692
15-Z-β-carotene	0.600
13-Z-β-carotene	0.596
9-Z-β-carotene	0.579
β-Cryptoxanthin	0.709
zeaxanthin	0.651
β-Carotene + β-Cryptoxanthin	0.582
β-branch proVA carotenoids	0.610
xanthophylls	0.706
cis-β-carotenes	0.624
proVitamin A	0.665
Total Carotenoids	0.687
Ratio1: β-branch/α-branch	0.736
Ratio2: βC/βCX	0.644
Ratio3: βC/(βCX + ZEA)	0.637
Ratio4: βCX/ZEA	0.841
Ratio5: (βC + βCX)/ZEA	0.733
Ratio6: βCX/proVA	0.821
Ratio7: βCX/TA	0.845

Narrow sense heritability was calculated for each carotenoid and derived trait. Calculating with three replicates.

Table 1.4 QTL found in the biparental population Hi27 x A272 based on LC

Trait	Chr.	Peak position cM	Support interval cM	Additive effect	Marker Nearest Peak	Peak LOD	PVE (%)	Candidate Gene
α-Carotene	5	63.8	57.8 - 66.8	0.07	PHM4647.8	7.47	9.72	
Lutein	1	185.7	174.9 - 194.2	0.32	PHM3034.3	3.95	3.89	*nced1*
	7	23.2	14.1 - 39.0	0.41	PZA00132.17	4.05	6.24	*zds1*
	8	39.5	19.9 - 78.5	0.33	pza02746.2	3.53	9.23	
	8	160	159.4 - 162.7	0.42	PZA01623.3	5.00	13.2	
	8	170.5	163.2 - 172	0.44	PHM4512.38	5.59	14.5	
	9	58.9	55.0 - 68.8	-0.43	PHM13183.12	4.81	4.84	
β-Carotene	1	178.9	170.9 - 187.9	0.25	PHM3034.3	4.38	3.94	*nced1*
	5	63.8	61.8 - 67.2	0.37	PHM4647.8	5.49	6.12	
	5	75.6	71.6 - 79.8	0.42	PZA00547.6	4.37	7.19	*lcyB*
	8	78.6	77.2 - 80.8	0.35	PZA00045.1	5.97	6.8	
	8	87.4	84.7 - 91.9	0.37	PHM4134.8	6.57	7.74	
15-Z-β-carotene	5	70.8	69.1 - 71.6	0.14	PZA02862.10	7.50	14	
	7	19.6	14.1 - 23.2	-0.11	PZA00153.3	4.83	5.71	*zds1*
13-Z-β-carotene	1	179.2	173.5 - 189.5	0.05	PHM3034.3	3.93	2.68	*nced1*
	5	66.8	64.0 – 68.0	0.10	PHM4165.14	7.79	10.4	
	5	74.6	68 - 80.1	0.10	PZA00547.6	8.08	13.1	
	8	88.9	78.2 - 92.8	0.08	PHM4134.8	4.40	4.24	
	8	116.6	106.3 - 126	0.09	PZA00196.2	5.60	9.25	
9-Z-β-carotene	7	2.8	0.0 - 7.4	0.11	PHM4135.15	5.66	7.48	
	7	11.7	9.7 - 15.7	0.10	PZA01613.1	4.18	7.99	
	8	78.6	78.5 - 82.1	0.12	PZA00045.1	5.55	10.8	
	8	87.4	84.5 - 91.2	0.13	PHM4134.8	7.30	12.1	
β-Cryptoxanthin	6	0.0	0.0 - 2.7	0.33	PHM15961.13	4.29	5.57	
	7	32.3	20.8 - 42.6	-0.25	PZA00084.2	4.30	11.9	
Zeaxanthin	5	63.8	62.6 - 64.3	1.92	PHM4647.8	10.70	16.9	
	5	69.1	68.9 - 70.8	2.05	PHM1870.20	12.38	18.2	*lcyB, nced8, cyp13, ggh1, hds1, mecs2*
	7	31.8	23.7 - 42	-1.12	PZA00084.2	4.49	10.2	
	8	72	57.5 - 75.9	1.30	PZA03005.19	5.35	2.33	*lcyE*
	8	78.6	77.7 - 80.8	1.34	PZA00045.1	5.03	3.66	
	10	123	119.3 - 129.2	1.05	PZA00337.3	3.75	2.77	
β-Carotene + β-Cryptoxanthin	5	63.8	63.0 - 66.6	0.80	PHM4647.8	10.05	12.1	
	5	70.8	68.3 - 71.6	0.80	PZA02862.10	10.21	13.2	*lcyB*
	9	2.0	0.0 - 13.8	0.55	PZA03062.7	4.56	6.7	
β-branch proVA carotenoids	1	177.9	172.1 - 179.2	0.54	PZA00381.3	3.97	3.14	*nced1*
	5	63.8	63.0 - 64.2	1.09	PHM4647.8	11.00	14.8	
	5	75.6	71.6 - 79.6	1.16	PZA00547.6	11.42	17.1	*lcyB*
Xanthophylls	1	177.9	172.2 - 187.7	1.05	PZA00381.3	3.51	3.22	*nced1*
	5	63.8	62.7 - 64.3	1.90	PHM4647.8	7.77	12.0	
	5	69.1	68 - 70.5	2.16	PHM1870.20	10.25	15.0	*lcyB, nced8, cyp13, ggh1, hds1, mecs2*
	8	63.8	53.1 - 74.9	1.90	PZA03005.19	6.56	6.45	*lycE*
	8	78.6	77.2 - 81	1.66	PZA00045.1	5.53	4.93	
cis-β-carotenes	1	177.9	174.2 - 190.4	0.17	PZA00381.3	4.39	2.81	*nced1*
	5	70.1	67.8 - 78.7	0.26	PZA02862.10	6.48	7.1	*lcyB*
	5	137.8	125.5 - 149	0.15	PHM3612.19	3.54	1.14	
	8	116.6	105.6 - 128.3	0.28	PZA02249.4	6.30	12	

Note: Chr, Chromosome; PEV, phenotypic variance explained

Table 1.4 Cont. QTL found in the biparental population Hi27 x A272 based on LC

Trait	Chr.	Peak position cM	Support interval cM	Additive effect	Marker Nearest Peak	Peak LOD	PVE (%)	Candidate Gene
Provitamin A	5	63.8	62.3 - 67.2	0.61	PHM4647.8	5.27	5.66	
	8	78.2	77.1 - 82.1	0.50	PZA00045.1	4.16	4.19	
	8	87.4	83.3 - 91.8	0.52	PHM4134.8	4.77	5.71	
	8	92.8	92.8 - 96.8	0.44	PHM4552.6	3.63	4.8	
	5	63.8	62.3 - 67.2	0.61	PHM4647.8	5.27	5.66	
Total Carotenoids	1	179.2	175.1 - 191.7	1.51	PHM3034.3	4.26	2.67	*nced1*
	5	69.1	67.8 - 70.7	3.27	PHM1870.20	12.05	16.6	*lcyB, nced8, cyp13, ggh1, hds1, mecs2*
	8	79.1	79 - 81.6	2.44	PZA00045.1	5.94	5.35	
	8	87.4	84.1 - 90.6	2.53	PHM4134.8	6.67	6.1	
	8	92.8	92.8 - 96.9	2.07	PHM5235.8	4.59	5.54	
Ratio1	5	70.8	70.11 – 76.0	0.58	PZA02862.10	13.82	15.77	*lcyB*
	7	23.2	22.6 – 23.2	-0.55	PZA00132.17	11.99	13.49	
Ratio2	1	163.4	156.4 - 170.9	0.07	PHM12706.14	3.98	3.79	
	6	1.5	0 - 4.5	-0.10	PHM15961.13	5.24	6.63	
	6	8.6	5.8 - 9.8	-0.11	PZA00578.1	6.50	6.07	
	8	73.5	65.3 - 77.2	0.10	PZA03005.19	5.35	7.52	*lcyE*
	8	78.6	77.8 - 80.3	0.11	PZA00045.1	6.05	8.85	
	8	88.4	85 - 92.2	0.11	PHM4134.8	6.41	10.3	
Ratio3	2	55.1	53.6 - 65.2	0.01	PHM3668.12	3.48	3.28	
	2	109.5	103.4 - 116.7	-0.02	PZA02080.1	8.91	10.9	*hyd4, nced4, ccd7, crti2*
	6	14.1	10.6 - 28.3	-0.01	PHM12904.7	4.62	3.54	*y1*
	10	60.5	57.5 - 70.1	-0.01	PZA01456.2	4.55	13.9	*crtRB1*
Ratio4	1	62.4	56.2 - 67.4	-0.01	PHM4597.14	4.05	1.86	
	1	70.8	69.7 - 78.2	-0.01	PZA00274.7	5.20	3.63	
	6	0	0 - 2.1	0.02	PHM15961.13	6.91	9.45	
	6	8.6	4.5 - 9.8	0.01	PZA00578.1	5.00	3.39	
	8	60.3	50.3 - 70.1	-0.02	PZA03005.19	6.73	7.2	*lycE*
	8	78.6	77.9 - 79.9	-0.02	PZA00045.1	6.50	6.39	
Ratio5	2	103	99.5 - 109.7	-0.03	PZA02080.1	6.94	8.36	*hyd4, nced4, ccd7, crti2*
	8	54.3	39.9 - 70.7	-0.02	pza02746.2	3.85	4.53	*lcyE*
Ratio6	1	164.4	161.8 - 168.6	-0.03	PHM12706.14	5.46	4.96	*nced1*
	3	55.3	55.1 - 60.2	0.02	PHM4955.12	4.90	1.9	
	3	66	64.9 - 68.8	0.02	pza01396.1	4.98	2.79	
	6	0	0 - 3.1	0.04	PHM15961.13	8.33	7.35	
	6	9.3	7.4 - 9.8	0.04	PHM8327.18	8.95	6.31	*y1*
	7	22.2	17.6 - 26.8	-0.02	PZA00153.3	3.61	7.03	*zds1*
	7	39.3	26.8 - 51	-0.02	PZA00084.2	3.86	9.5	
	8	89.9	87.4 - 92.3	-0.04	PHM4134.8	7.26	8.33	
Ratio7	1	60.2	56.6 - 67	-0.01	PHM4597.14	4.46	2.39	
	1	70.8	69.8 - 77	-0.01	PZA00274.7	5.26	3.55	
	6	0	0 - 2.2	0.01	PHM15961.13	7.80	10.4	
	6	9.3	5 - 9.8	0.01	PHM8327.18	6.27	4.29	*y1*
	8	89.9	87.4 - 92.8	-0.01	PHM4134.8	5.42	7.13	
	8	103.6	93.2 - 115.3	-0.01	PZA00196.2	6.38	9.65	

Note: Chr, Chromosome; PEV, phenotypic variance explained

Table 1.5 QTL found in the biparental population Hi27 x A272 based on visual score and colorimeter from the Owens, 2015.

Trait	Chr.	Peak position cM	Support interval cM	Additive effect	Marker Nearest Peak	Peak LOD	PVE (%)	Candidate Gene
Visual 2013	1	147.2	125.1-162.9	0.33	PHM1438.34	4.19	8.68	
	1	170.6	167.3-177.9	0.28	PHM12693.8	4.39	6.01	*nced1*
	6	43.3	27.6-52.3	-0.43	PZA02048.2	8.25	11.88	
	9	69.4	61.5-72.6	0.43	PHM7916.4	8.12	12.65	
	10	51.9	42.9-63.4	-0.44	PHM15868.56	8.31	13.81	
h	5	57.3	50.2-62.5	-2.30	PHM9009.13	7.95	12.56	*nced9*
	5	67.8	67.2-69	-2.57	PHM2769.43	11.08	15.03	
	5	74.6	71.6-79.6	-2.24	PZA00547.6	7.78	12.01	*lcyB*
	6	36.3	25-46.7	2.70	PZA02048.2	7.65	15.05	
	6	56.3	46.7-64.6	2.41	PHM12794.47	8.86	13.51	
	9	66.4	61.5-72.3	-3.01	PHM2278.86	12.11	19.14	
	10	64.6	45.5-73.1	1.35	PZA02519.7	4.38	5.50	
h99	5	57.3	50.2-62.5	-2.12	PHM9009.13	8.23	12.97	*nced9*
	5	67.8	67.2-69	-2.39	PHM2769.43	11.69	15.87	
	5	74.6	71.6-79.6	-2.08	PZA00547.6	8.23	12.94	*lcyB*
	6	36.3	25.1-46.7	2.40	PZA02048.2	7.53	14.38	
	6	56.3	46.7-64.7	2.18	PHM12794.47	8.88	13.43	
	9	66.4	61.4-72.1	-2.77	PHM2278.86	12.59	19.81	
	10	64.6	45.5-73.1	1.21	PZA02519.7	4.31	5.41	
L*	6	39.3	25.9-46.7	1.97	PZA02048.2	7.49	11.19	
	6	57.3	46.7-66.3	1.82	PHM12794.47	8.13	11.61	
	9	48.1	42.1-53	-2.53	PZA02861.12	9.92	15.98	*dxs3*
	9	65.4	61.3-70.3	-3.30	PHM2278.86	17.78	27.70	
	10	63.9	48.9-64.1	1.89	PZA02519.7	9.93	13.61	
a*	1	209.3	197.8-224.2	0.50	PZA02957.5	4.54	7.01	*ao1,ao3,ao4*
	9	60.4	62.1-72.4	1.08	PHM2278.86	12.98	24.32	
b*	3	40.6	39.4-42.8	1.94	PHM5502.31	4.48	5.74	*dxr1,ggh2, nced3*
	3	51.6	48-54.6	2.43	PZA00920.1	7.54	9.24	
	3	58.3	55.3-61.5	2.16	PHM4955.12	5.88	7.96	
	4	160.2	136.7-170.2	-1.58	PHM5665.10	3.73	4.11	*cyp15,hyd7*
	5	67.8	67.2-70	-3.04	PHM2769.43	11.70	15.03	
	5	73.6	70.8-79.6	-2.74	PZA00547.6	9.68	13.13	*lcyB*
	6	58.3	52.2-66.2	3.33	PHM12794.47	12.57	17.74	
	6	69.3	66.3-77.1	3.20	PZA01342.2	10.98	14.85	*dxs1,ippi3*
	9	67.4	60.4-72.4	-2.38	PHM7916.4	6.25	8.70	
	10	25.6	22.5-29.3	-2.23	PZA00587.3	4.03	4.30	
L	6	40.3	25.9-46.7	1.82	PZA02048.2	7.55	11.27	
	6	57.3	46.7-66.3	1.76	PHM12794.47	8.16	11.62	
	9	48.1	42.1-53	-2.28	PZA02861.12	10.07	16.26	*dxs3*
	9	65.4	61.3-70.2	-3.00	PHM2278.86	18.13	28.19	
	10	63.9	49.3-64.1	1.80	PZA02519.7	10.07	13.74	
a	1	209.3	197.1-221.2	0.47	PZA02957.5	6.00	9.76	*ao1,ao3,ao4*
	3	50.2	46.7-61.3	0.42	PZA00920.1	3.83	6.51	

Note: Chr, Chromosome; PEV, phenotypic variance explained

Table 1.5 Cont. QTL found in the biparental population Hi27 x A272 based on visual score and colorimeter from the Owens, 2015.

Trait	Chr.	Peak position cM	Support interval cM	Additive effect	Marker Nearest Peak	Peak LOD	PVE (%)	Candidate Gene
b	5	67.8	65-69.6	-1.38	PHM2769.43	10.38	12.46	
	6	40.3	26.6-46.7	1.37	PZA02048.2	8.44	12.09	
	6	57.3	48.8-66.3	1.50	PHM12794.47	11.02	14.58	
	9	47.1	42.1-51.7	-1.39	PZA02861.12	8.66	11.71	*dxs3*
	9	65.4	60.7-72.1	-1.69	PHM2278.86	12.96	17.93	
	10	25.6	22.7-29	-1.35	PZA00587.3	4.50	5.03	
	10	43.9	42.6-44.9	2.01	PZA01456.2	10.39	13.44	*crtRB1*
	10	51.9	45.9-61.1	1.84	PHM15868.56	9.15	15.31	
L99	6	40.3	25.9-46.7	1.86	PZA02048.2	7.42	11.08	
	6	56.3	46.7-66.3	1.75	PHM12794.47	8.08	11.56	
	9	48.1	42.1-53	-2.47	PZA02861.12	9.73	15.64	*dxs3*
	9	65.4	61.2-70.5	-3.12	PHM2278.86	17.34	27.08	
	10	63.9	48.3-64.1	1.82	PZA02519.7	9.77	13.78	
a99	1	17.2	12.2-34.7	0.22	PZA00175.2	3.59	5.86	*hyd8*
	1	209.3	198.7-222.8	0.24	PZA02957.5	5.84	9.19	*ao1,ao3,ao4*
	3	52.6	44.1-55.3	0.31	PZA00920.1	6.73	10.60	*nced3*
	9	67.4	65.8-72.6	0.32	PHM7916.4	7.11	12.11	
	9	82.6	72.6-93.7	0.36	PHM7916.4	7.84	14.74	
b99	3	52.6	42.8-55.3	0.63	PZA00920.1	4.25	5.21	*nced3*
	5	67.8	67.2-69.2	-1.10	PHM2769.43	13.76	17.95	
	6	57.3	47.9-66	1.06	PHM12794.47	11.29	16.35	
	9	67.4	66.2-75.8	-1.03	PHM7916.4	9.91	14.13	
	10	25.6	22.7-29.3	-0.77	PZA00587.3	3.82	5.23	
	10	54.9	38.3-64.3	0.64	PHM15868.56	4.03	5.90	*crtRB1*
C99	3	40.6	39.5-42.8	0.55	PHM5502.31	7.06	9.46	*dxr1,ggh2, nced3*
	3	53.6	48.9-55.3	0.66	PHM4955.12	10.47	13.74	
	5	66.3	65.2-68	-0.63	PHM4165.14	10.64	14.04	
	6	59.3	50.6-66.3	0.79	PHM12794.47	14.44	21.60	

Note: Chr, Chromosome; PEV, phenotypic variance explained

Table 1.6 Candidate genes for the biparental population Hi27 by A272

Chromosome	Candidate Gene	RefGen V2 ORF Start	RefGen V2 ORF Stop	Maize Reference genome V2
1	*hyd8– hydroxylase 8*	6,353,416	6,354,652	GRMZM5G826824
1	*ao1 - aldehyde oxidase 1*	286,448,581	286,456,365	GRMZM2G141535
1	*ao3 - aldehyde oxidase 1*	286,358,278	286,366,211	GRMZM2G019799
1	*ao4 - aldehyde oxidase 1*	286,506,118	286,513,080	GRMZM2G141473
1	*nced1 – 9-cis-epoxycarotenoid dioxygenase 1*	250,892,567	250,895,242	GRMZM2G014392
2	*hyd4– carotene hydroxylase*	15,865,938	15,868,219	GRMZM2G164318
2	*nced4 – 9-cis-epoxycarotenoid dioxygenase 4*	234,574,835	234,576,854	GRMZM2G408158
2	*ccd7 – carotenoid cleavage dioxygenase 7*	19,458,968	19,461,625	GRMZM2G158657
2	*crti2 - carotene isomerase 2*	226,366,352	226,371,341	GRMZM2G106531
3	*dxr1 - deoxy xylulose reductoisomerase 1*	30,226,804	30,233,358	GRMZM2G056975
3	*ggh2 - geranyl hydrogenase 2*	40,062,008	40,064,270	GRMZM2G419111
3	*nced3 – 9-cis-epoxycarotenoid dioxygenase 3*	87,344,791	87,346,554	GRMZM5G858784
4	*hyd7– hydroxylase 7*	236,023,117	236,025,051	GRMZM2G163683
4	*cyp15 - cytochrome P450 15*	235,724,340	235,728,875	GRMZM2G010221
5	*cyp13 - cytochrome P450 13*	215,827,224	215,831,730	GRMZM5G837869
5	*lcyB – lycopene beta cyclase*	100,700,176	100,702,026	GRMZM5G849107
5	*nced8 – 9-cis-epoxycarotenoid dioxygenase 8*	200,687,176	200,689,579	GRMZM2G150363
5	*nced9 – 9-cis-epoxycarotenoid dioxygenase 9*	16,850,172	16,851,977	GRMZM5G838285
5	*ggh1 – geranylgeranyl hydrogenase 1*	206,890,298	206,892,838	GRMZM2G105644
5	*hds1 – Hydroxymethylbutenyl diphosphate synthase 1*	182,124,005	182,130,631	GRMZM2G137409
5	*mecs2 - methyl erythritol cyclodiphosphate synthase 2*	196,279,295	196,281,037	AC209374.4_FG002
6	*dxs1 - deoxy xylulose phosphate synthase 1*	146,378,393	146,382,661	GRMZM2G137151
6	*ippi3 - isopentenyl pyrophosphate isomerase 3*	147,131,116	147,136,679	GRMZM2G133082
6	*y1-phytoeno synthase*	82,017,148	82,021,007	GRMZM2G300348
7	*ggps2 - geranylgeranyl pyrophosphate synthase 2*	160,531,537	160,533,586	GRMZM2G102550
7	*zds1 - zeta-carotene desaturase1*	17,470,585	17,479,020	GRMZM2G454952
8	*lcyE – lycopene epsilon cyclase1*	138,882,594	138,889,812	GRMZM2G012966
9	*dxs3 - deoxy xylulose phosphate synthase 3*	14,077,852	14,081,075	GRMZM2G493395
10	*crtRB1 – Beta-carotene hydroxylase*	136,057,100	136,060,219	GRMZM2G098676

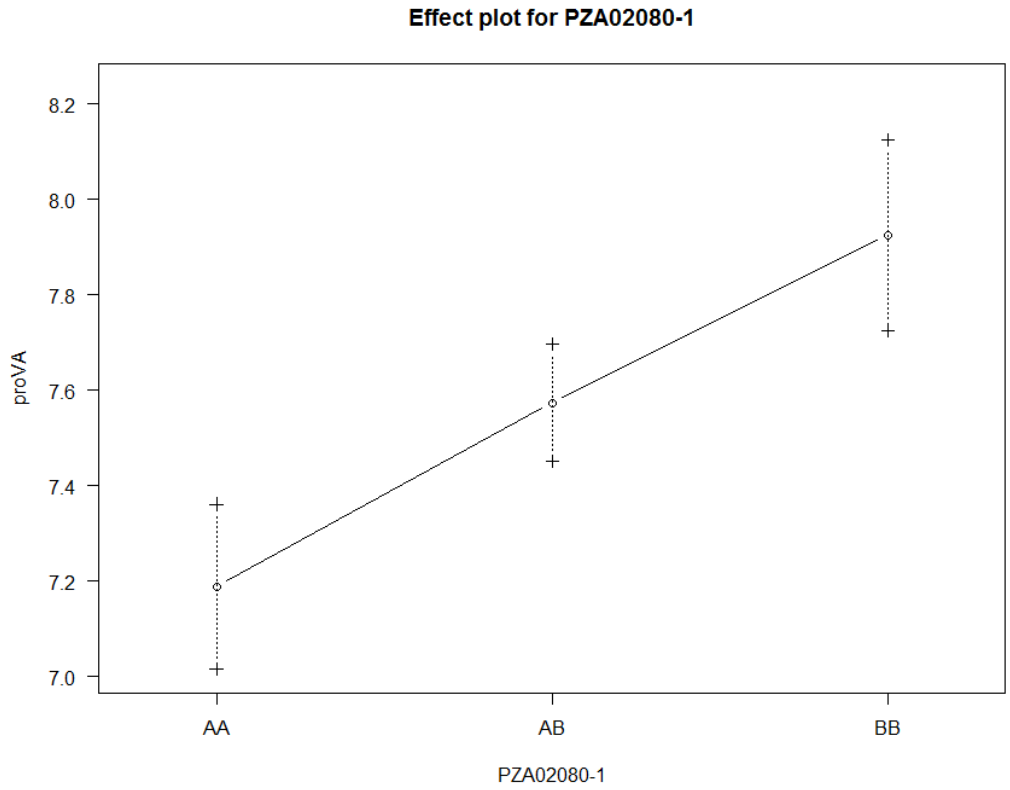


Figure 1.6 Effect plot for concentration of proVA by the SNP PZA02080-1. AA and BB are parental alleles, the AB correspond to the heterozygous allele. The values for proVA for AA is lower compared with the BB for the marker SNP PZA02080-1 which is the peak SNP for the QTL that underlie the gene *hyd4*

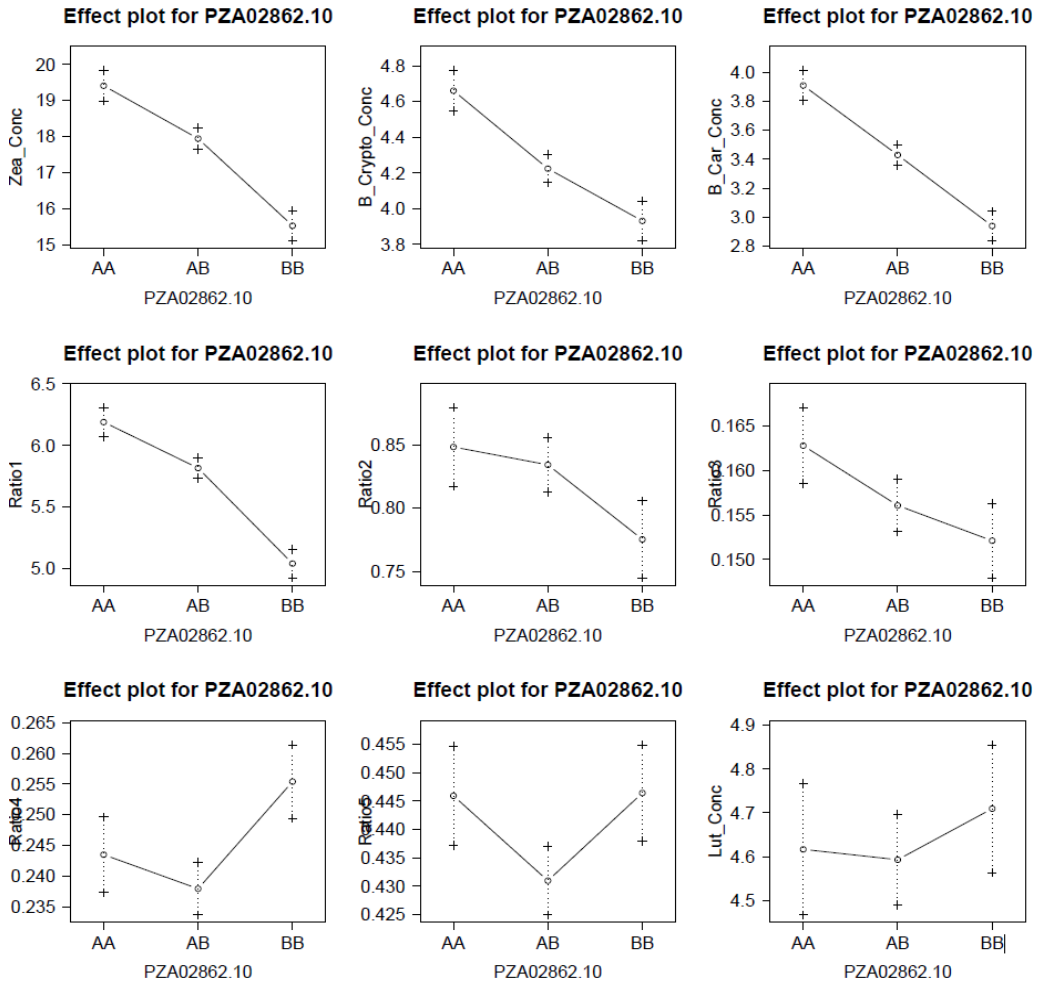


Figure 1.7 Effect plot for the marker SNP PZA02862.10 (peak marker closest to lcyB). The traits analyzed were zeaxanthin, β-cryptoxanthin, β-carotene, Ratio1, Ratio2, Ratio3, Ratio4, Ratio5 and lutein.

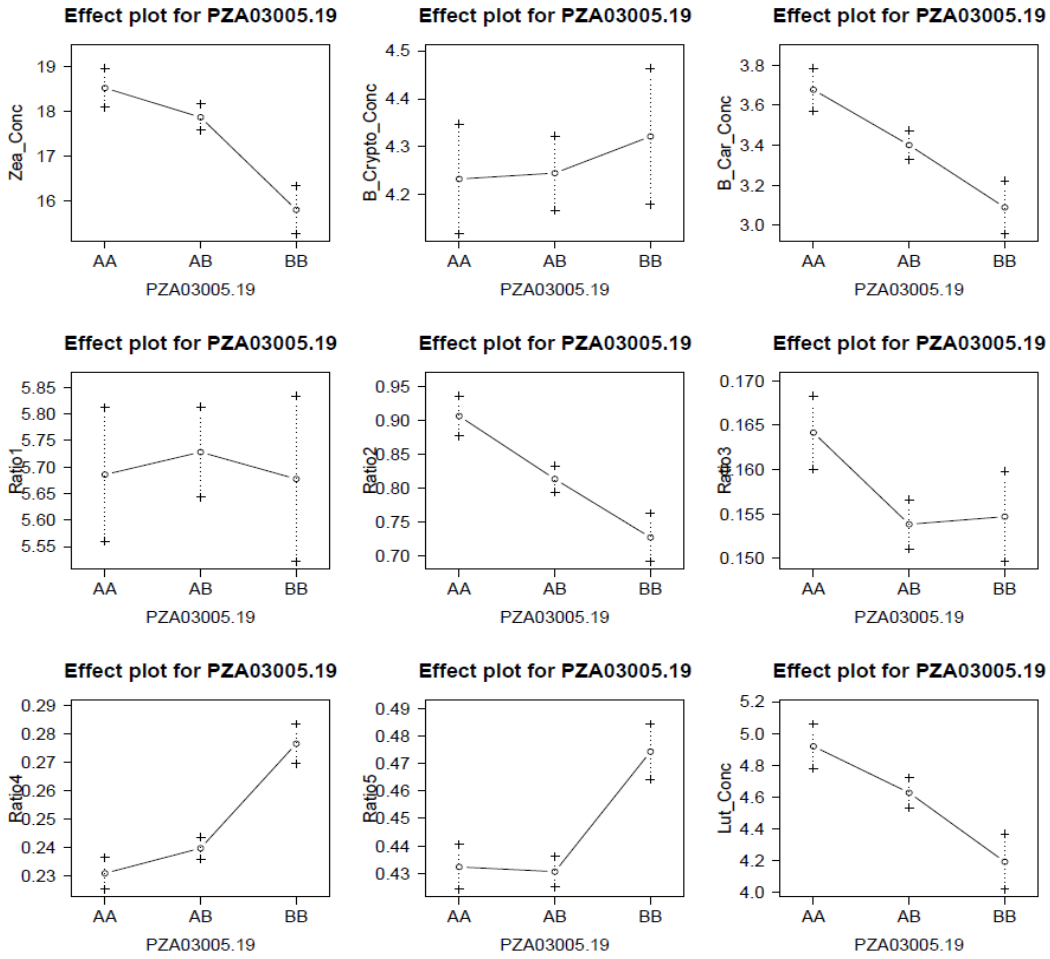


Figure 1.8 Effect plot for the marker SNP PZA03005.19 (peak marker closest to lcyE). The traits analyzed were zeaxanthin, β-cryptoxanthin, β-carotene, Ratio1, Ratio2, Ratio3, Ratio4, Ratio5 and lutein.

A

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	0.11639294	0.01454912	4.20	0.0001
Error	198	0.68520778	0.00346065		
Corrected Total	206	0.80160072			

R-Square	Coeff Var	Root MSE	Ratio2 Mean
0.145201	13.41086	0.058827	0.438654

Source	DF	Anova SS	Mean Square	F Value	Pr > F
PZA02080_1	2	0.07857127	0.03928564	11.35	<.0001
PZA00578_1	2	0.00960321	0.00480161	1.39	0.2521
PZA02080_*PZA00578_1	4	0.02821846	0.00705461	2.04	0.0905

B

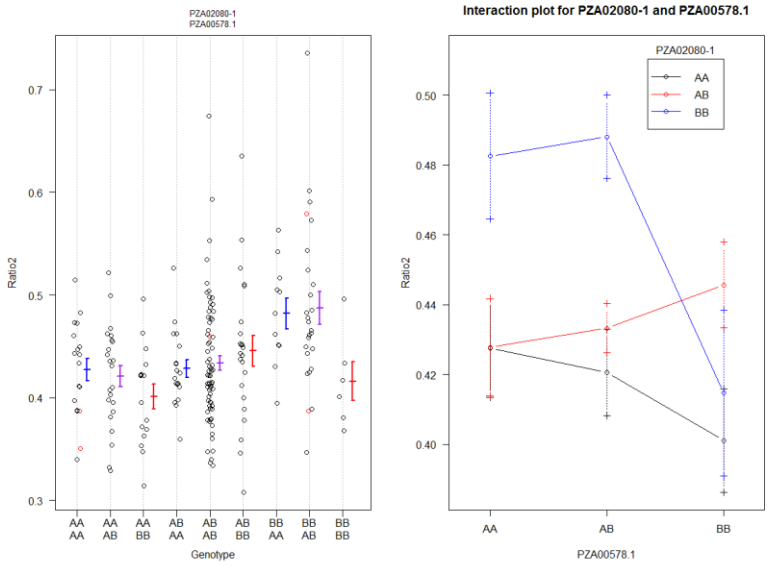


Figure 1.9. Interaction between peak snp for *hyd4* (PZA02080.1) and *y1* (PZA00578.1). A) The SNP PZA02080.1 is significant and the interaction with PZA00578.1 is slightly significant B) Plot effect for *hyd4* vs *y1* a trend is not clear to observe

APPENDIX

APPENDIX

Signature of Selection for darker orange kernels

Plant Material

A total of fifteen double haploid (DH) lines and four inbred lines were used for this study. A synthetic population was developed from the four public inbred lines: KI3, KI11, KI43 and KIU2007. These inbreeds were developed by Kasetsart University in Thailand and they are all derived from the Suwan breeding population (Sriwatanapongse et al. 1993). A series of F1 crosses among the four inbreds were made at the University of Illinois at Urbana-Champaign in 2003. The F1 hybrids were crossed in a winter nursery in Hawaii in 2003-2004. The four-way hybrids were random-mated in Urbana in summer 2004. In Urbana in 2005, reciprocal crosses (IxII and IIxI) were made by bulking pollen from the top ear. The IxII cross resulted in darker ears. Of the 24 IxII ears, the top 9 were selected and of the 24 IIxI ears, the top 10 were selected. The reciprocal crosses were combined to make a balanced bulk, talking 160 kernels per ear (one ear with only 143 kernels). Remnant seed of the selected ears was combined in an unbalanced bulk. Unselected ears were also combined in an unbalanced bulk. The resulted ears were ranked by intensity of orange color, and the top ears were used to create an isolated open pollinated population (cycle 0). Recurrent selection based on darker orange ears was done in the field in subsequent cycles.

From cycle 1 to 8 the synthetic population was planted in isolation at the University of Illinois at Urbana-Champaign and at Purdue University, and the same procedure for selection for darker orange ears was carried out in each cycle. The best ears from cycle 8 were sent to AgReliant for double haploid (DH) procedure. Two replicates were planted of the four progenitors and fifteen DH lines in summer 2015 at the Agronomy Center for Research and Education (ACRE) with plot rows of 4.5 m long and 13 plants and one replicate was planted in a greenhouse three pots per entry. Plants were self-pollinated and the ears were bulk harvested, shelled and stored at -80 °C until phenotyping.

Carotenoid extraction and quantification by LC

Carotenoid extraction was performed as described by Ortiz et al. (2016). For each bulk harvest 10 g of maize grain were taken at 45 days after pollination. The 10 g were ground to a fine powder (0.5 mm) in a Foss CT 1093 Cyclotec. A double acetone extraction and a methyl tert-buthyl ether extraction were used. After extraction, the samples were resuspended in 1 mL of methanol and 1 mL of ethyl acetate. Only a 10µL sample was injected into a Hewlett-Packard 1090 HPLC with a diode array detector and a YMC C30 column (150 x 20 mm). Two different mobile phases were employed: mobile phase A (Methanol: Ammonium Acetate, pH = 4.6, 98:2) and mobile phase B (Ethyl acetate) with a gradient: 0 – 6 min 85% A to 15% B; 6 – 8 min 20% A to 80% B; 8 – 12 min 100% B and 14 – 15 min 85% A to 15% B. An internal standard (β-apo-8'-carotenal) was used and eight carotenoids were measured: lutein (LUT), zeaxanthin (ZX), β-cryptoxanthin (βCX), β-carotene (βC), α-carotene (αC) and isomers15-Z-β-carotene, 13-Z-β-carotene and 9-Z-β-carotene. From these values proVA carotenoids and total carotenoids (TC) were

calculated. The content of each carotenoid, proVA and TC were reported in µg • g-1 of dry weight (DW).

DNA extraction, whole genome sequencing and SNP calling pipeline

Genomic DNA was extracted from ten lyophilized leaves using the cetyl(trimethyl) ammonium bromide (CTAB) protocol (Warburton 2005). The whole genome sequencing was done at Macrogen, Inc. with a library target size of 250 bp. Pair-end (PE) sequencing was performed in the Illumina HiSeq X. Figure 1 shows all the step followed for the SNP calling variants with a similar workflow as described by (Altmann et al. 2012). Quality control was assessed with the FastQC (v 0.11.2) for all the samples. The quality trimming was done using the software cutadapt (version 1.9.1) to remove adapter in both the 3' and the 5' ends with less than Phred33 score of 30. The Deconseq tool (version 0.4.3) was used to detect any contamination in the data. After quality trimming the reads were mapped against the bowtie2-indexed B73 RefGen_v3 using Bowtie2 package (version 2.2.9) with number of allowed mismatches as 1. Duplicates read were removed with the Picard tools MarkDuplicates (version 2.3.0).

For the SNP and InDel calling we used both the GATK UnifiedGenotyper and the samtools mpileup. Before calling the SNP with the UnifiedGenotyper, we performed a local realignment across all reads. RealignerTargetCreator determined suspicious interval which are likely in need of realignment and then running the realigner over those interval using InderRealigner. The base quality score recalibration was done by applying machine learning to model the errors empirically and adjust the quality scores for the variant calling algorithm. Positions at which both SNP and InDels were called, were filtered out

(McKenna et al. 2010; DePristo et al. 2011; Auwera et al. 2013). For samtools mpileup the base alignment quality was turned off because it can cause a real SNP to be missed. The positions that have an average mapping quality as 0 were filtered (Li 2011).

SnpEff was used (after variant calling and filtering) to annotate and classify genetic polymorphisms based on their effects on annotated genes, such as synonymous or non-synonymous SNPs start codon gains or losses, stop codon gains or losses; or based on their genomic region location, such as intronic 5' UTR, 3' UTR, upstream, downstream or the intergenic regions (Cingolani et al. 2012). We used the common findings between the GATK and the Samtool resulting in a combining dataset of 4246 SNP for our subsequent analysis.

Signature of selection and haplotype blocks

From the WGS a set of 59 candidate genes from the isoprenoid and the carotenoid biosynthetic pathway were studied. From this set of genes a subset of genes were chosen for detecting signature of selection (SS). The variant called (SNP) were subject to a second quality control. This process involved a) Only a SNP that satisfied phred-scaled confidence value (default is 30) were chosen and b) heterozygous variant or indels were removed. Allele frequencies were estimated for both the parental and the DH lines. A FST was used to identify significant SNPs (p-value = 0.05). We corrected for multiple comparisons through false discovery rate (FDR) using the Storey method (Storey 2002). This method is a more traditional statistical procedure yet powerful compared with the Benjamini-Hochberg. This method fixes the rejecting regions and estimate α for multiple testing.

Under this method an analogous of p-value is calculated, that is the q-value and it gives a hypothesis testing error for each observation considering the positive FDR (pFDR).

Haplotype block (HB) were defined by the Four gametes rules (FGR), this rule assumes that only 3 gametes are observed from 4 possible combination based on the two alleles (Wang et al. 2002)

References

Altmann A, Weber P, Bader D, et al (2012) A beginners guide to SNP calling from high-throughput DNA-sequencing data. Hum Genet 131:1541–1554.

Auwera GA, Carneiro MO, Hartl C, et al (2013) From FastQ data to high- confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinforma 10–11.

Cingolani P, Platts A, Wang LL, et al (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6:80–92.

DePristo MA, Banks E, Poplin R, et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993.

McKenna A, Hanna M, Banks E, et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303.

Ortiz D, Rocheford T, Ferruzzi MG (2016) Influence of Temperature and Humidity on the Stability of Carotenoids in Biofortified Maize (Zea mays L.) Genotypes during Controlled Postharvest Storage. J Agric Food Chem 64:2727–2736.

Sriwatanapongse S, Jinahyon S, Vasal SK (1993) Suwan-1: Maize from Thailand to the world. CIMMYT

Storey JD (2002) A direct approach to false discovery rates. J R Stat Soc Ser B (Statistical Methodol 64:479–498.

Wang N, Akey JM, Zhang K, et al (2002) Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. Am J Hum Genet 71:1227–1234.

Warburton M (2005) Laboratory Protocols CIMMYT Applied Molecular Genetics Laboratory. CIMMYT: International Maize and Wheat Improvement Center

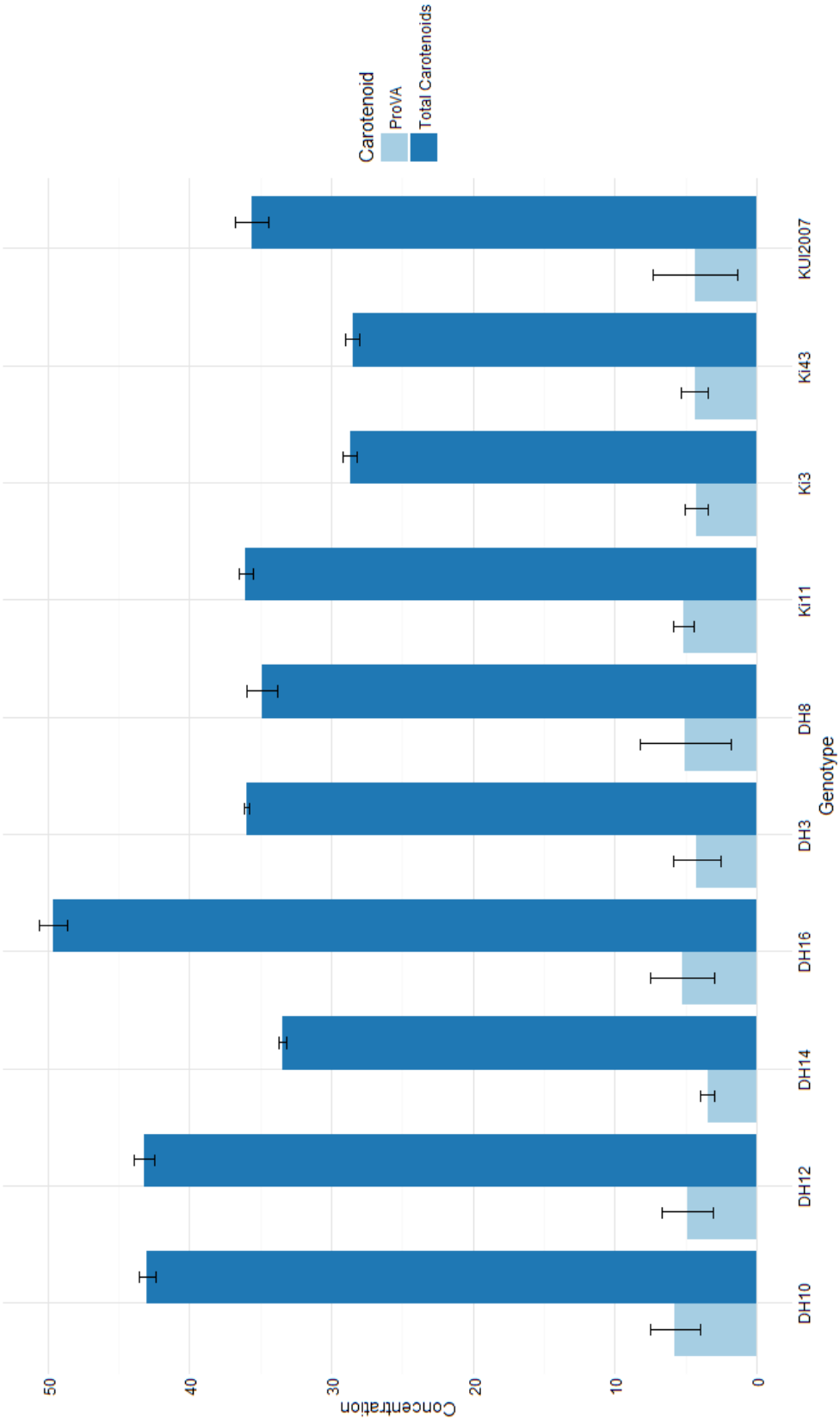
Appendix Table 1. Carotenoid content in the four parents and the fifteen DH lines.

Genotype	Carotenoids																			
	LUT		ZEA		βCX		αC		βC		15ZβC		13ZβC		9ZβC		proVA		Total Carotenoids	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
KI3	3.95	0.35	17.83	0.93	2.79	0.34	0.32	0.25	1.64	0.44	0.96	0.38	0.58	0.09	0.62	0.06	4.27	0.81	28.68	1.38
KI11	9.58	1.03	18.74	1.51	2.15	0.07	0.49	0.35	2.59	0.52	0.87	0.19	0.84	0.08	0.77	0.35	5.15	0.87	36.03	1.25
KI43	8.78	0.30	13.31	0.20	1.98	0.62	0.41	0.15	2.32	0.36	0.54	0.11	0.54	0.09	0.62	0.15	4.37	0.83	28.51	1.65
KIU2007	6.57	1.06	22.36	1.80	2.39	0.34	0.40	0.32	1.98	1.43	0.56	0.16	0.56	0.23	0.82	0.43	4.34	2.05	35.65	5.13
DH1	3.64	0.68	26.21	2.36	2.76	0.45	0.43	0.11	2.93	0.82	0.88	0.27	0.65	0.21	0.59	0.50	5.58	1.43	38.09	4.49
DH2	3.47	0.20	25.91	1.60	2.53	0.25	0.33	0.09	2.77	0.25	0.75	0.15	0.60	0.07	0.73	0.09	5.24	0.45	37.10	0.82
DH3	4.49	0.04	25.40	2.43	1.46	0.22	0.41	0.09	2.39	0.14	0.48	0.24	0.60	0.02	0.74	0.22	4.23	0.28	35.97	2.87
DH4	3.55	0.25	21.88	1.91	2.52	0.36	0.45	0.50	2.19	0.77	0.82	0.36	0.63	0.43	0.97	0.75	4.89	1.96	33.01	4.48
DH5	4.71	0.68	28.84	3.48	2.10	0.53	0.41	0.03	2.65	0.46	1.13	0.29	0.78	0.30	0.73	0.07	5.22	0.87	41.35	5.41
DH6	3.68	0.67	26.21	1.19	4.15	0.81	0.36	0.12	3.64	0.52	1.12	0.29	0.92	0.13	0.75	0.16	7.28	1.29	40.81	3.91
DH7	3.39	0.18	23.39	0.68	4.42	0.51	0.39	0.08	2.48	0.05	0.98	0.27	0.72	0.15	0.80	0.18	6.14	0.13	36.58	0.55
DH8	3.23	0.53	24.10	2.33	2.48	0.87	0.34	0.09	2.57	0.95	0.79	0.37	0.57	0.19	0.78	0.26	5.05	1.83	34.87	5.57
DH9	3.90	0.43	28.64	3.07	1.55	0.49	0.42	0.44	2.96	1.32	0.77	0.51	0.95	0.46	1.01	0.56	5.31	2.51	40.19	6.99
DH10	4.97	0.31	29.18	3.18	2.81	0.26	0.39	0.16	2.67	0.63	1.12	0.55	0.92	0.17	0.93	0.08	5.76	1.01	43.00	3.01
DH12	5.22	0.72	30.88	0.92	1.91	0.23	0.29	0.23	2.67	1.03	0.69	0.24	0.72	0.14	0.77	0.16	4.86	1.26	43.16	3.11
DH13	6.51	0.69	27.47	2.09	2.11	0.43	0.42	0.24	3.28	0.87	0.82	0.39	0.91	0.27	0.92	0.28	5.87	1.57	42.44	4.76
DH14	5.43	0.19	22.87	0.79	1.36	0.12	0.36	0.16	1.82	0.47	0.55	0.12	0.52	0.02	0.53	0.10	3.48	0.52	33.44	0.90
DH15	3.51	0.50	28.37	1.38	1.43	0.23	0.25	0.19	2.52	1.33	0.67	0.16	0.58	0.21	0.72	0.19	4.34	1.78	38.05	3.49
DH16	4.61	0.68	37.14	0.79	2.35	0.54	0.40	0.20	2.66	1.03	0.80	0.29	0.75	0.28	0.88	0.18	5.26	1.76	49.60	3.93

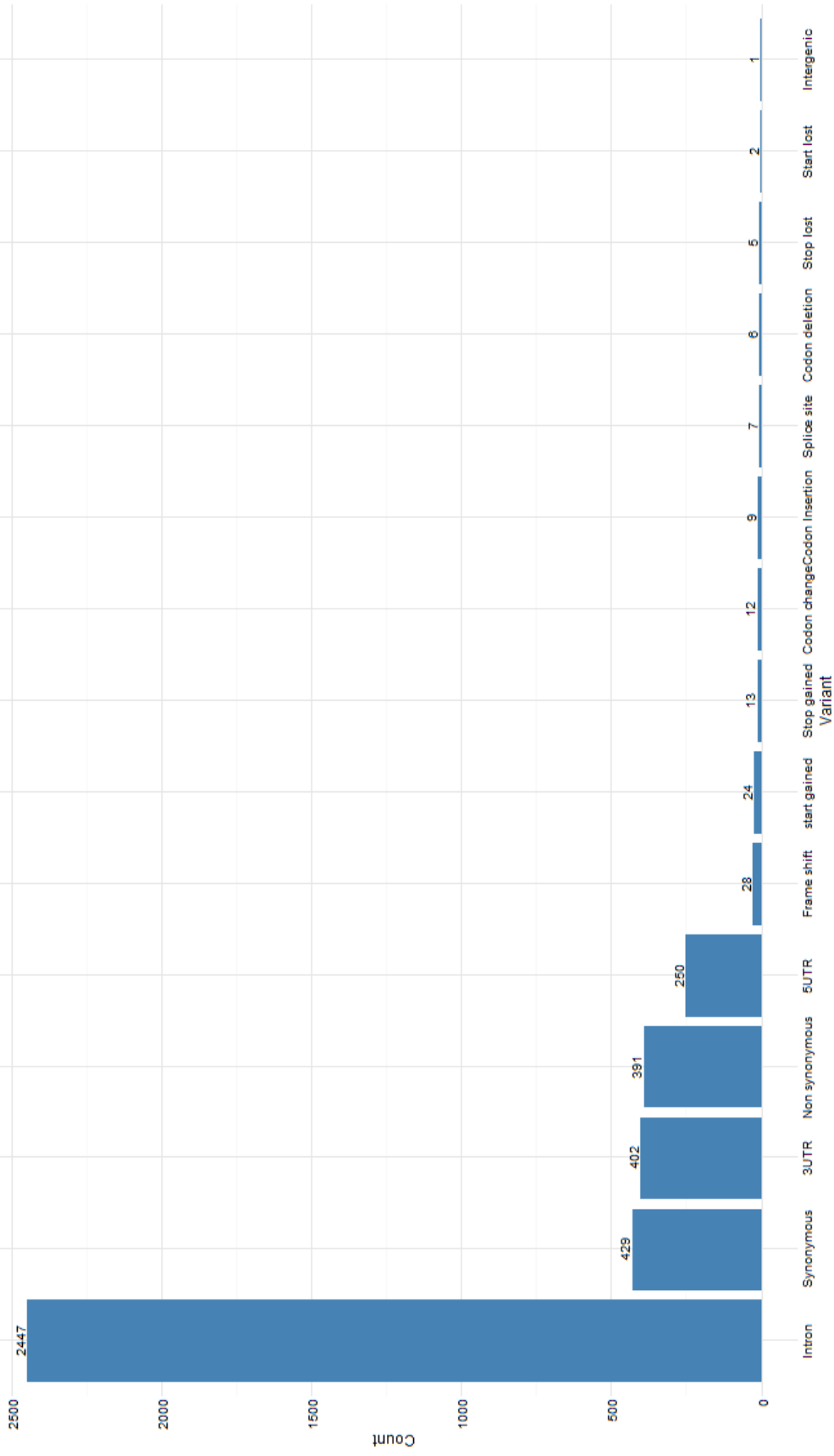
Eight carotenoids where measure Lutein (LUT), zeaxanthin (ZEA), β-cryptoxanthin (βCX), α-carotene (αC), β-carotene (βC) and the isomer, 9-Z-β-carotene (9ZβC), 13-Z-β-carotene (13ZβC) and 15-Z-β-carotene (15ZβC). Two values where calculated the provitamina A (proVA) and the total carotenoids (TC).

Appendix Table 2. Candidate genes analyzed in the isoprenoid and the carotenoid biosynthetic pathway

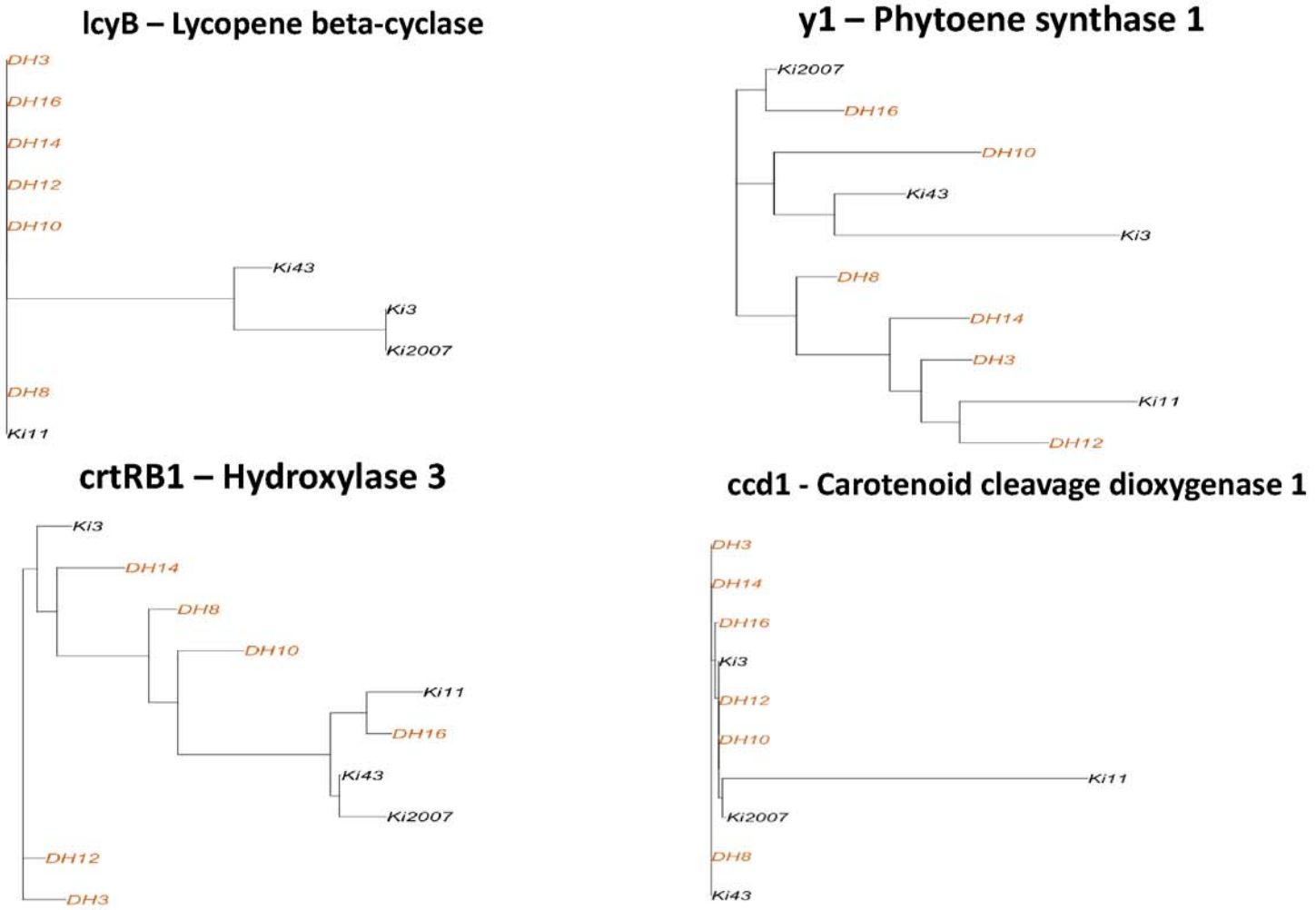
Candidate Gene	Annotated gene function	Chr.	Variant Detected					TOTAL*
			5' UTR	Intron	Non-synonymous coding	Synonymous coding	3' UTR	
dxs2	1-deoxy-D-xylulose 5-phosphate synthase	7	0	1	0	4	12	22
hyd4	Beta-carotene hydroxylase 4 (non-heme dioxygenase type)	2	8	28	4	13	5	60
zds1	Zeta-carotene desaturase 1	7	4	104	4	3	3	119
zep1	Zeaxanthin epoxidase 1	2	16	195	21	10	37	280
y1	Phytoene synthase 1	6	2	1	1	0	8	12
lut1	CYP97A3, Cytochrome P450 epsilon-ring hydroxylase	1	0	26	0	0	6	32
lycB	Lycopene beta-cyclase	5	0	0	2	5	4	11
crtRB1	Beta-carotene hydroxylase 1 (non-heme dioxygenase type)	10	1	14	2	1	12	34
lycE	Lycopene epsilon-cyclase	8	18	0	5	8	1	184
ccd1	Carotenoid cleavage dioxygenase1	9	2	37	6	5	7	58
vp14	9-cis-epoxycarotenoid dioxygenase1	1	4	0	4	8	8	24
	Total		55	406	49	57	103	836
	Percent Type		6.58%	48.56%	5.86%	6.82%	12.32%	100.00%



Appendix Figure 1. Carotenoid concentration for the DH lines and the parents KI3, KI11, KI43 and KIU2007. The total carotenoids are calculated based on the eight carotenoids measured in this study, whereas the proVA content is calculated based on those carotenoids with vitamin A activity.

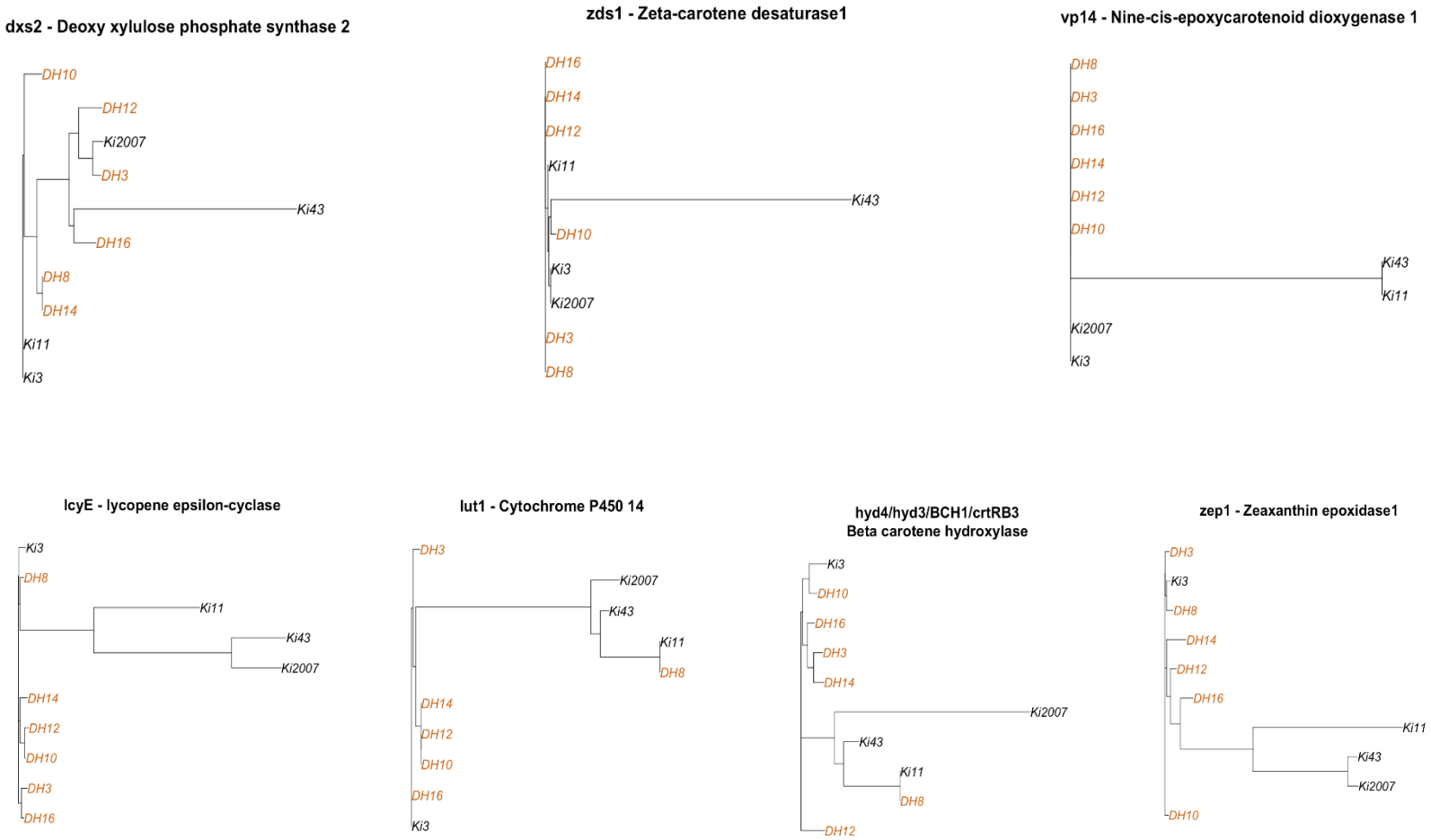


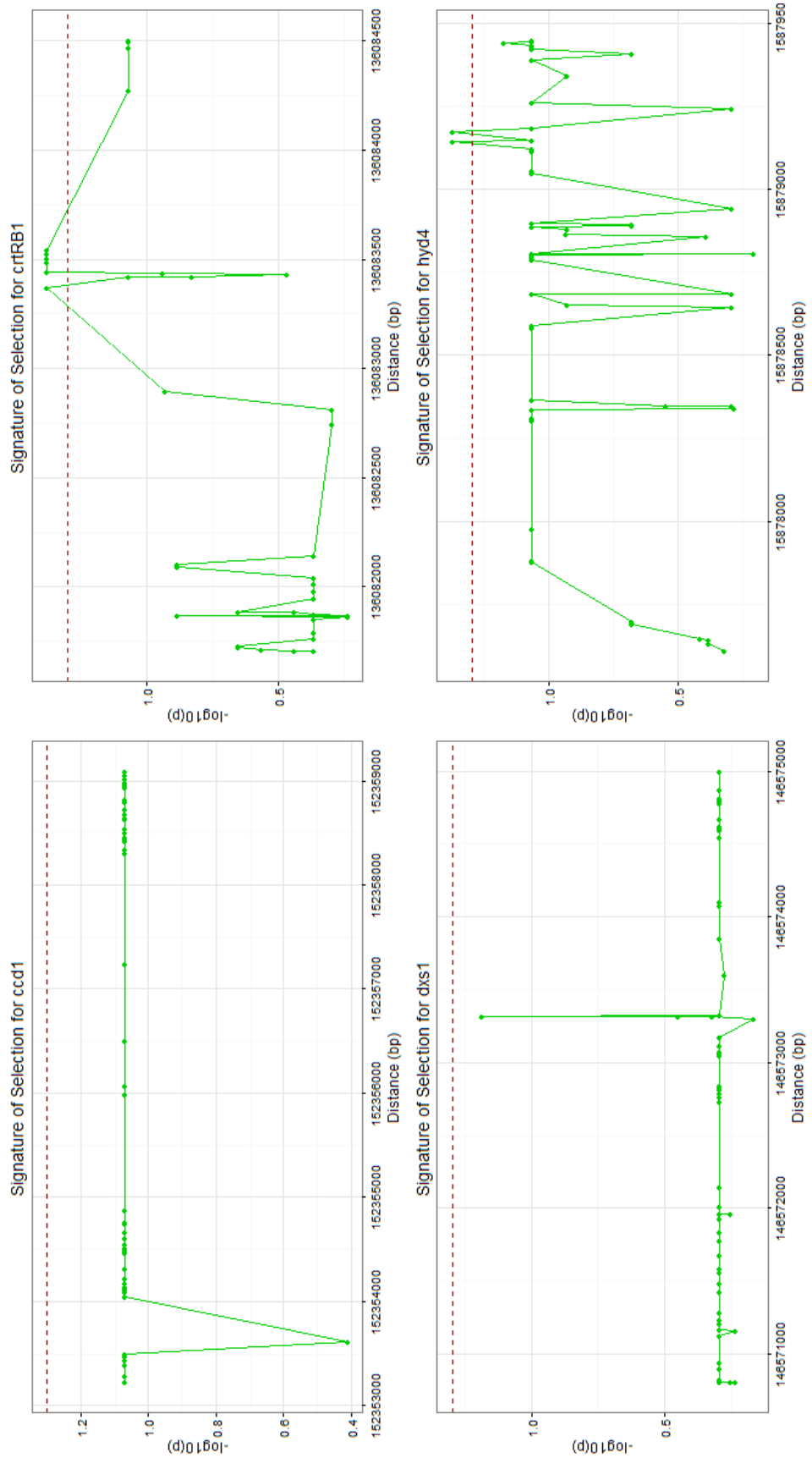
Appendix Figure 2. The number and density of variant called. The SnpEff software assigned the SNP and classified them in different categories. This results are before the filtering process



Appendix Figure 3. Strong selection for the parental line Ki11 for *lcyB*, while the converse occurs suggesting strong selection against the gene *ccd1* for this parent. Other genes display less distinct SNP-based patterns of association (*hyd3* and *y1*).

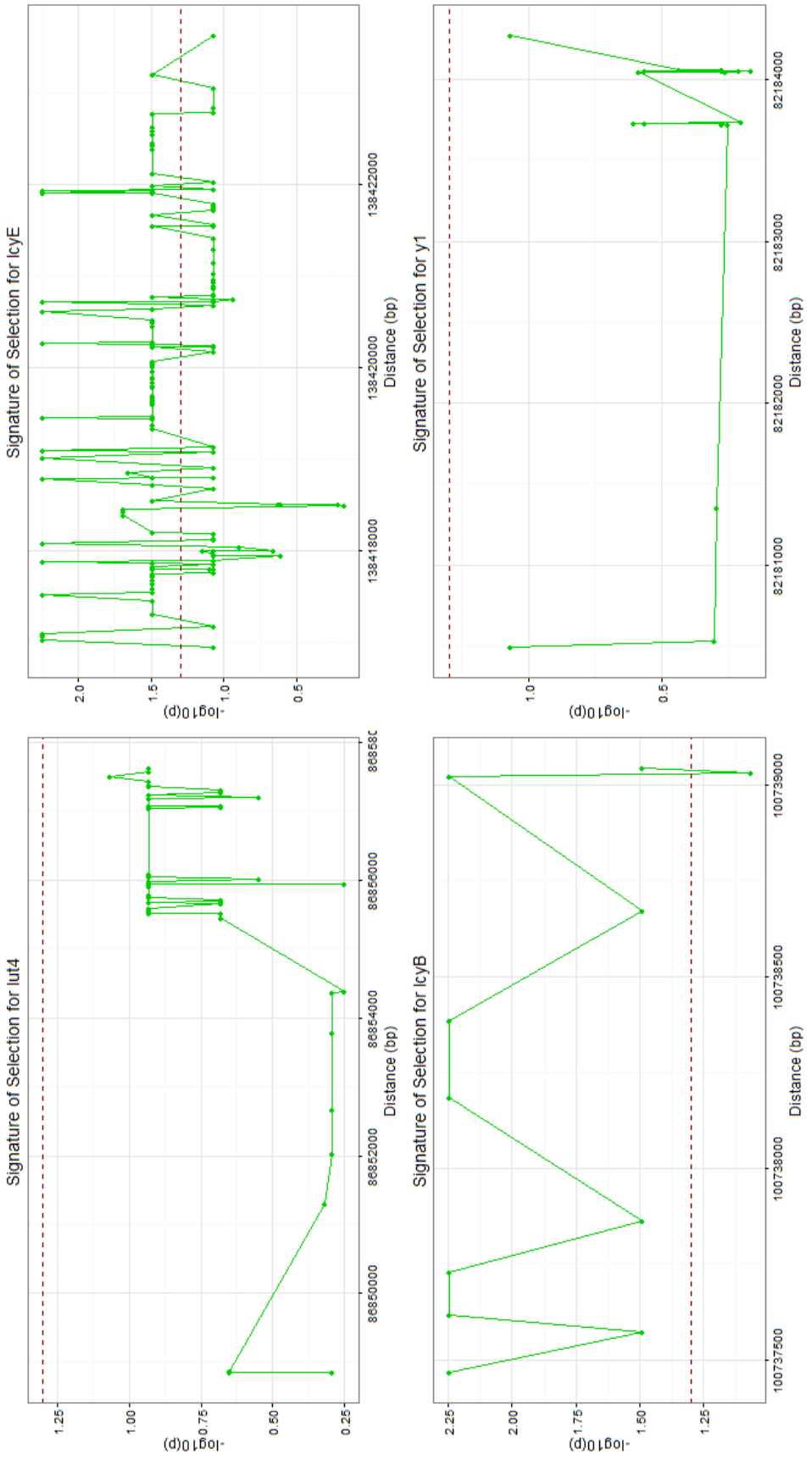
Appendix Figure 3. Continued



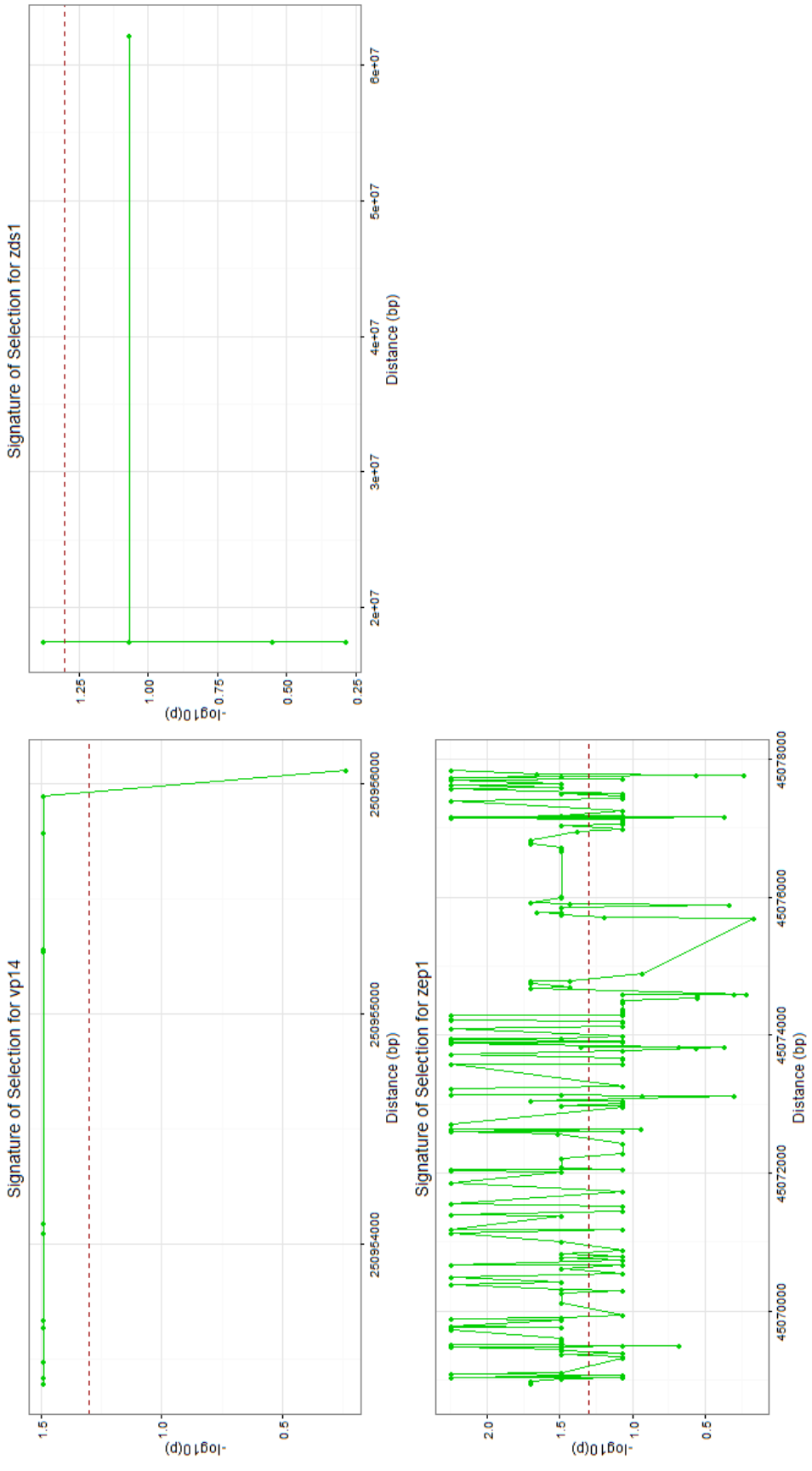


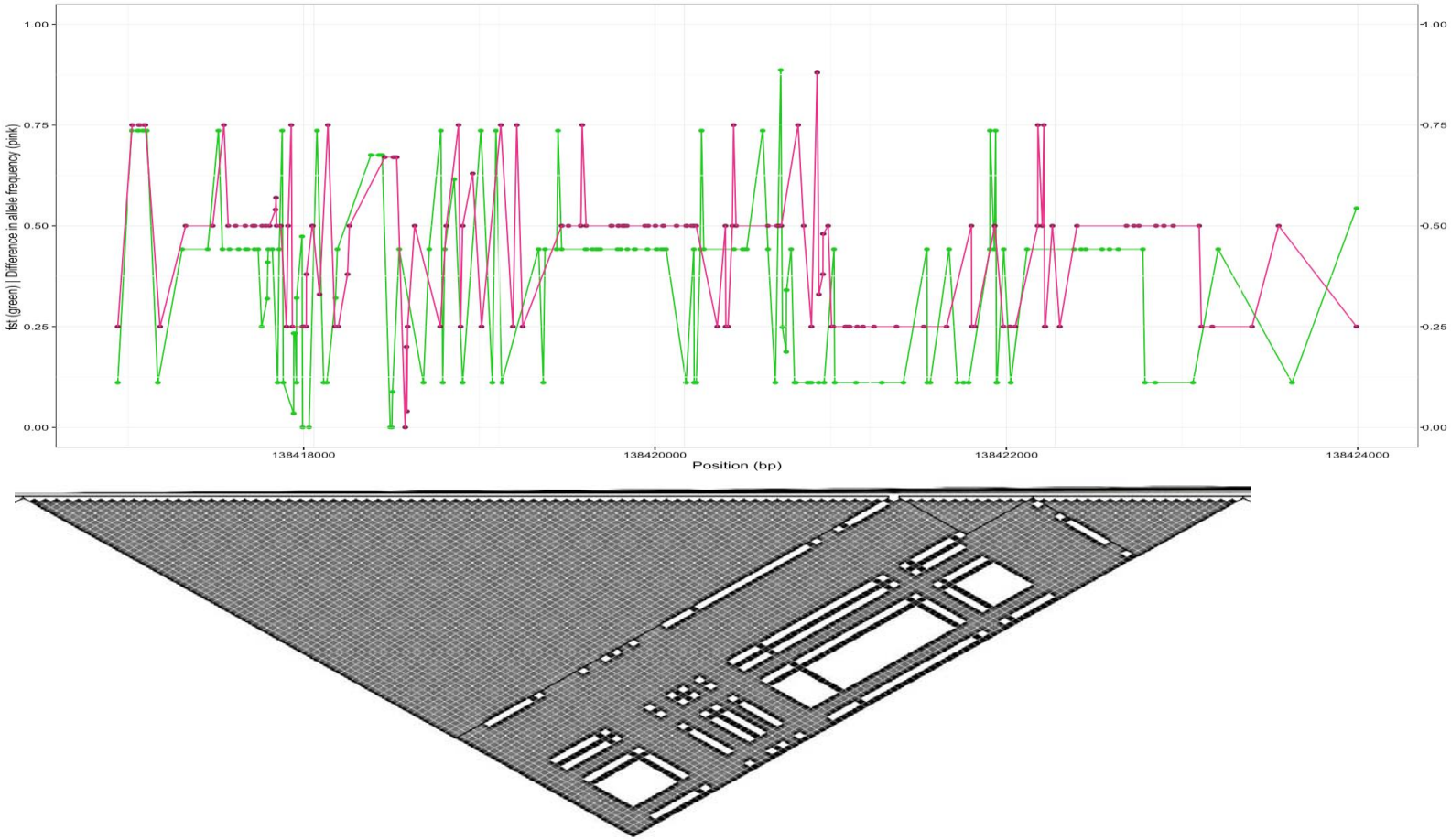
Appendix Figure 4. Signature of selection for a subset of genes. The SNP that are above the threshold are significant for the gene, in some genes the selection was across most of the SNP in other cases it wasn't significant.

Appendix Figure 4. Continued



Appendix Figure 4. Continued





Appendix Figure 5. Haplotype block for LcyE. The green plot represent the F_{ST} for LcyE whereas the pink is the difference in the allele frequencies between the parental and the DH lines. The haplotype block shows a solid black block evidence of selection for this gene